

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2003, 07:54:40 ; Search time 38 Seconds

(without alignments)
2287.245 Million cell updates/sec

Title: US-09-150-867-1

Perfect score: 14769
Sequence: 1 MSEDGAKVCVRVRLQRE.....QAEWMYEAKEKTAPECKTS 2954

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/2/1aa/PCITUS.COMB.pep:*
- 6: /cgn2_6/ptodata/2/1aa/backfilest.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1148.5	7.8	3248	1	US-08-353-700-1
2	1148.5	7.8	3248	5	PCT-US95-16216-1
3	1055.5	7.1	2482	1	US-08-328-254-6
4	1051	7.1	1388	4	US-09-572-191-2
5	1051	7.1	1388	4	US-09-723-262-2
6	1051	7.1	1388	4	US-09-723-219-2
7	953	6.5	1234	4	US-09-592-054-8
8	945.5	6.4	1232	4	US-09-592-054-2
9	906	6.1	1375	4	US-09-722-139-2
10	906	6.1	1375	4	US-09-721-832-2
11	906	6.1	1375	4	US-09-721-689-2
12	866	5.9	10182	4	US-09-134-001C-3159
13	841.5	5.7	1279	4	US-09-724-517-2
14	841.5	5.7	1279	4	US-09-641-807A-2
15	841.5	5.7	1279	4	US-09-723-096-2
16	769	5.2	3696	4	US-09-134-001C-5080
17	763	5.2	955	2	US-08-428-414A-3
18	759	5.1	955	1	US-08-006-676B-1
19	759	5.1	955	1	US-08-282-845-2
20	759	5.1	955	5	PCT-US94-00324-1
21	758.5	5.1	1066	4	US-09-541-782-8
22	758.5	5.1	1066	4	US-09-723-820-8
23	756.5	5.1	1637	4	US-09-718-692-2
24	756.5	5.1	1637	4	US-09-718-852-2
25	756.5	5.1	1637	4	US-09-718-815-2
26	748.5	5.1	2101	1	US-08-195-487-4
27	748.5	5.1	2101	5	PCT-US93-06160-4

28	744.5	5.0	2101	1	US-08-466-390-4	Sequence 4, Appl1
29	744.5	5.0	2101	1	US-08-470-950-4	Sequence 4, Appl1
30	744.5	5.0	2101	1	US-08-467-781-4	Sequence 4, Appl1
31	744.5	5.0	2101	2	US-08-483-924-4	Sequence 4, Appl1
32	744.5	5.0	2101	4	US-09-452-294-1	Sequence 1, Appl1
33	733	5.0	1073	4	US-09-541-782-6	Sequence 6, Appl1
34	733	5.0	1073	4	US-09-723-820-6	Sequence 6, Appl1
35	723	4.9	473	4	US-09-592-054-6	Sequence 6, Appl1
36	723	4.9	522	4	US-09-592-054-4	Sequence 4, Appl1
37	717.5	4.9	1898	1	US-08-056-200-94	Sequence 4, Appl1
38	717.5	4.9	1898	2	US-08-800-644-94	Sequence 94, Appl1
39	704	4.8	1184	4	US-09-541-782-2	Sequence 2, Appl1
40	704	4.8	1184	4	US-09-723-820-2	Sequence 2, Appl1
41	702	4.8	1057	4	US-09-541-782-10	Sequence 10, Appl1
42	702	4.8	1057	4	US-09-723-820-10	Sequence 10, Appl1
43	696	4.7	411	2	US-08-713-815A-4	Sequence 4, Appl1
44	696	4.7	441	2	US-08-713-815A-3	Sequence 3, Appl1
45	692	4.7	1053	4	US-09-724-519-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1
US-08-353-700-1Sequence 1, Application US/08353700
Patent No. 5599919

GENERAL INFORMATION:

APPLICANT: YEN, TIMOTHY J.

APPLICANT: RATNER, JEROME B.

TITLE OF INVENTION: NUCLEIC ACID ENCODING A

TITLE OF INVENTION: TRANSGENICALLY EXPRESSED KINETOCHORE PROTEIN,

TITLE OF INVENTION: AND METHODS OF USE

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESSES:

ADDRESSEE: DANN, DOREMAN, HERRELL AND SKILLMAN

STREET: 1601 MARKET STREET, SUITE 720

CITY: PHILADELPHIA

STATE: PA

COUNTRY: USA

ZIP: 19103-2307

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/353,700

FILING DATE: 09-DEC-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: REED, JANET E.

REGISTRATION NUMBER: 36,252

TELECOMMUNICATION INFORMATION:

TELEPHONE: (215) 563-4100

TELEPHONE: (215) 563-4044

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 3248 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: HUMAN

US-08-353-700-1

Query Match 7.8%, Score 1148.5, DB 1, Length 3248;
Best Local Similarity 20.3%, Pred. No. 2.3e-48;
Matches 715; Conservative 626; Mismatches 1180; Indels 995; Gaps 139;

QY 106 QAIQIE---VKIIOETIPNREFLLRVSYMETIYEWYKOLICDRKKPLPIREDENRNY 161
18 OKIOELEGOLDKIKKEKQOFOLDISLEAPQOKOTK---VENEKTEGMLKRENDRIAE 74
QY 162 VADITEELVAVPEHVIOMIKKGENRHYGETKMANDHSSRSHTIFRMIVESRDNNDPTNSE 221
Db 75 ICSELETKOKISHELO-VKESQYN--FOEGOLNSGKKQLEKLEQEL--KRCSELEERSQ 129
QY 222 NCDGAVVSHLN-----LVDLASERASOTGAEVRLKEGCINNSLFTLGQV--- 269
Db 130 QAAQSAADVS--LNPENPTOKIETTPPLTPSOYSGSKYEDLEKKNYKEERKRLAEAEKAL 188
QY 270 -IKLSLSCQAGGFINYDLSKILRLQNS-----LGNATVYICITTPPSFD 315
Db 189 QAKKASQTLFOATMNRHD--TARHQASSVPSWQOETPSPHLSNSQR-----TPIRRD 240
QY 316 -----ETLSILOFA-----STAKHVNTPHNEVLDDEALIKRYKEIILDK 357
Db 241 FSASYFSGELEVPTRSTLQIGKRDANSFPFGNSSPHLLDOL--KAQONELRKINELE 298
QY 358 KOLENLESSSE-----TKAQAMAKEHTOLLAETIKOLIKEREDRIWHNLNIYVASSQES 411
Db 299 LRIQHEKEMKGYNKFOEIOLOLEKAKVELIEKEKVLNCRDBLVFTTAQYDQASTKYT 358
QY 412 QODQRYVR-----KRYVTWAPGKIONSLHAGVSDFDMLSRLPGNF-----SKR 455
Db 359 ALBOKIKKLTLDLSCQONANESARCSLEPQIKKEKEKEFOELSRQOSFQTLDOECIQMK 418
QY 456 AKFS--DMSPEPIDSVCTEFSDFDALSMDSNGIDAEMNL-----ASKYTHKEKSLH 509
Db 419 ARLOEOLOAKNMHNVLQAEIDKLTSYKOOLENNLEEFKOKLRAEOAFQSOQKEKELR 478
QY 510 OSMIDFGOISDVQFHDSSKENOLOYLPKDGSMAECHKAFKEKITSILOQLOSKREEK 569
Db 479 RSMERKKNLILSHSOKARBEVCHLEAKLNKOCLOANS-----QNFMEEMAKMTSQ 533
QY 570 KELVQSFELKIALEEOISVYKAKLEMYTNSREHSINAEOVDYKEVYKREMSYVLGDSG 629
Db 534 ETMLRDLOEKINQOENSLYTEKILKAVA-----DLEKORDOSODILKRR-----EHN 580
QY 630 YNASSDLODSSVQKRLSSSHDECIERHMKLEOKIYDLEFETELNK-----KSENDK-- 683
Db 581 IEQULDKISTEKEKSKALISA-----LELKATYEBELKEKILFSCWKSENEKL 630
QY 684 -QKSEODFMES---IOLC-----EATMAEKANALELALMRDNFNIT 723
Db 631 TQMESEKENLQSKINHLETCLKQOIKSHEYNERVRLTEMREMLSVETIRMLHNVLDKS 690
QY 724 LENETLKREIADLERS-----LKENQETNEFELIEKETQOKHEKOLTHE 767
Db 691 VEVEFOKLAUVELQOKAFESDQKQKEIENMCKLTSQLTQGVEDL-----EHKLOLILS- 743
QY 768 IGLSKLIYENAMNYONLEEDLETKTLKBOEIOU----- 803
Db 744 ---NEIMDKORCY--QDLHAETESLRDLIKSKDASLVINEDHQRSLAFDOQPMHNSFA 798
QY 804 -----AELRRKADN-----LOKVRNFDLSVSM-----GDSKICEETIFQ 838
Db 799 NITGGGSMPSRSECRLEADQSPKNSAILQNRVDSLEFSLESCQKMSNDILOKQCEELVQ 858
QY 839 LKOSISDAVATVYRDAQKECSFLRSLENLEKEMEDTSMNYN-----QKEKAASLTF 888
Db 859 IKGEITE--ENLMAKADMOHOSFYAETSORISKLOEDTSAHONVAVETLSALEKKEKELOLL 916
QY 889 EKOLETEKSNYKME-----ADLOKELO-----SAFNINYNLN 921
Db 917 NDKVETELBQAEIOPLKSNHLLDLSLKEIQLSETLSLEKEMSSITILNKRIETELQEN 976
QY 922 GLL-----AGRVPRDLSRV-----ELEKKVSEFSKOLE----- 950
Db 977 GYLKEINASLNOEKMLLQKSSSFANYIDERKESISELSDYQKOKELILLORCEETGNAY 1036
QY 951 -----KALEEKNA-----LENEVYCLSEYKFLPNEVECLKNOIKASSEIML-LKQSE 998

Db 1037 EDLSQKYKNAQKNKLEBOLINECTSLCENR--KNELBOLKEAFKQHOEFLTKIAFAEE 1094
QY 999 HSASITISKOELTIMQ-----EJQSE-----QILOLDETHTQNSVY---OOTE 1096
Db 1095 RNQONMLTELETVOOALRSEMTDNONNSKSEAGLKOETITLTKLEQONKQKEVNDLQONE 1134
QY 1037 EGYLEKKKHHND--DLFEKTYRKNSEADLLREMNILKQTE-----SVE 1078
Db 1155 QLMKVKYKTHQONLESEPTIRNSVYKERESERNQCNFKPQMDLEVKETISDYSYNAQVQLE 1214
QY 1079 VAIADTKHELEETIRDKOELLHEKKYFFQAMQOTIPPTPLSLSPPSYLVENSQDPLEI 1138
Db 1215 AMLRNKELIKQESKEKECLOHE-----LOTI-----RGLDETSLQDMQSOEISGL 1261
QY 1139 NDY-----HNLALMTERNNIMVCLTERNSLKEQVIDLNLQOLOSQASTE- 1185
Db 1262 KCEIDAERKYISGPELSTSONDANALQCSLOTTMKNKNE---LEKICEILOAKYEL 1317
QY 1186 -----KSDLOKPKODLEGEVYKLLLEMLLKG----- 1212
Db 1318 VPELNDRSBECITATRKMAEYVGKILNEVYKILNDGSLHGLVEDIPGEGFERQOPNEQH 1377
QY 1213 -----HTDSQLSIEKQLENLEVEYTEKQTLQLOEMKNITITERNELOTFED 1258
Db 1378 PVSIALPDESNSYEHILT---LSQEVQMPHAELOEFELSLQSBHKLHDHCOMSKSKSE 1434
QY 1259 LKAHDSLKODLSENIQSIETQODELRAAOEELREKQOLVDSFRQOLDCSVGISPPNDH 1318
Db 1435 LQTYVDSIK--AENVLVS---TNLRNFGDLVKEMQO--GLEBLVPSLSSQVPPDS 1485
QY 1319 AVANQEVKSIGEYVNSIOSEMLRGEDELOTSICALYSELE-----LMA 1362
Db 1486 SLVS-----SLGD--SSEYRALLE-----QDGMDSLNLNLEGAVSANQCVDEYFCSLOT 1533
QY 1363 HVSXVEGENEITRKNLNGLEKEITLQKSE--ESEVLKSMLENLEKDNKRLK----- 1410
Db 1534 YVDSIKRANLYVLTNLNRNFGDLVKEMQOGLLEGVLVPSLSSQVPPDSLSLGDSSFYR 1593
QY 1411 ---EQAEYS-----SKENQFSLEBYV-----SGSQKINDEI---E 1440
Db 1594 ALLEQTDMSLNLNLEGAVSANQCVDEYFCSLOEBENLTRKETPAPAKVGELESICE 1653
QY 1441 VYKQOLKAABERLE---IDRDYFELVQYANLVVE--GKLEPPLQADHE--EDSIDRS 1493
Db 1654 VYROSLEKLEKESQGYMKETIOELBOLLSBERQELDLCKROYLSENDQONOKLTSYV 1713
QY 1494 EEMEIKVLEKLEKERNQYILE---RLOEKLELSNKLLETLQKEMETSVLLKDDLOQKLES 1549
Db 1714 LEMSKILAAKKQOEOULSLEVARLOGLDILSSR-----SLGIDTPEADIOG 1762
QY 1550 LLSENITLKENIDITTLKHNHDQAOLOK-----TQOELOL--AKNULATAASD 1594
Db 1763 ---RNSQDISKEHTSETTERPKNDVHOICDKDAQOIDLNLTEKITERGAVKAPRG 1815
QY 1595 NCPITQOKET-----SADCVHPL-----EKKILLTEELHOK 1626
Db 1816 ECSEGEQSPDNIYEPGDEKTOQSSECSISELSPGPNALVPMDLGNEDLHNLQVARKET 1875
QY 1627 TNEOEKILHEKNELEQAOVELKCEVEHLMKSMIESKSLSELOHEKHDTQOOLA----- 1681
Db 1876 SNEMLRLLHYIEDRR-----KVESILNEMKELDSKL-----HIOEVOIOMTKYIEAC 1921
QY 1682 --LKQOMQVYQOEKKELOQTHEHTLAVDHLKENIEL--GLNFRNENQ--OKTKREOCLLN 1736
Db 1922 IELERKIYGLKKNESDISEKLEYFSCDHQELQOEVETSEGINSDLEHNAKSSREDIGDN 1981
QY 1737 ENK-----ELEQSOHRLQCE-----IEELMSLKDKESALTEKESQKY 1776
Db 1982 VAKYNDKSKKEPFLDVENELISIRSEKASISHEALYTLADLEEVQYTEKLCLEKMDENKQY 2041
QY 1777 I-NLNOEMVYMLMEELKNSQRTVIAER---DQLODLRESVMSIET----- 1821

```

Db 2042 IVCLEELSVTSEKNOQLGELDTMSKTTALDQLSEKKEKTKQ--ELFSEHCECHICQV 2100
Qy 1822 -ODDURKQOALQOQKQOELTQOISVQOEKISLEENOMLVNATVKTLEERDDLOS 1880
Db 2101 AAEKVEKTELLQTLSDVSELKDKTHLOEKLOQSLQSOALSLTKCLEMOQLOMKE 2160
Qy 1881 KOHLESEITLSLSKEKEFE-----ALEQEKDADARTIDITEKISNI----- 1926
Db 2161 KELLKESSELOARLESSEYKELNYSKALEALYKEREALALSTOEYHOLRGIEKL 2220
Qy 1927 -----EEOQLQATNKEKTLYERESLIQOEKOLALNTEHLRETLKSDALGMEQER 1979
Db 2221 RVRIADEKKQHLIAEKLEKEREENDST--KDKVENLERELQMESENEELVI--LDAEN 2275
Qy 1980 DEANAKVALATPKMSLEEQINENTYTLKEGE-----GEKETYLOQPSQOQSSQME 2032
Db 2276 SKA-----EVELTKQLEEMARSLKIFELDLVLRSEKENITKQIOEQOGOLSELD 2326
Qy 2033 ELRESLTKRDLQLEAEKEISEAT-----NEIKNLTKATISSL--EEREIQ--NASILN 2081
Db 2327 KILSFKSLLEKBOAEIQIKESKTVAMLONLKELNEAVAAALCGOELKATQESID 2386
Qy 2082 EAVSERENLRHSKQOVLSELE---QSLTLKSDHAFQSKREKDEAVNKIASIAEBI 2136
Db 2387 PPIEBHQLRNIEKLRARLEADEKKQCVLQOLK-----ESENHADLLKGVENLEREL 2441
Qy 2137 KILTEMD-----EPRDSK---ESLOEQSHLSEELCTYKTELQMKQKQEDINNKLAERV 2189
Db 2442 ELARTNOHAALEANSQGEVETLAKIEGMSQSLRGELDVVLTIRSEKENTINLEQKQ 2501
Qy 2190 KEVDLLQHLSSLEKQOLOQIOMELNREKLYELCEKMDIMEKISVRLKMONPEQOED 2249
Db 2502 ERISL-----EITNSSEN-----ILOEQEYVQMK 2530
Qy 2250 DVAERMDILESNOEIQIEMEKISAVYSEQHTLS--SLSELOKETBAKHCMLNIKE 2306
Db 2531 KSSTAMENIQ---OLKEINERVALAHNDQEACKQESQVE-----CL-----E 2575
Qy 2307 SLSSFLSSFGSLQTEHNKLNQLOTLNKKVUVYRTAAVKEDHSLINDYEDDLAAEOKR 2366
Db 2576 LEKADLOGLDBAKANNYVLOSVAGLIOE-----VEGKQKLEKKEE 2619
Qy 2367 HDELRLQCLEQHGKRSKDSASEELK---CEIEFLNELFKRANITQVODDFSEYO 2422
Db 2620 ISRLKNOIQDOEQVSKLSQVGEHQIMKEQNLRLNLTVELEOKTIOVLOSNAALSDTL 2679
Qy 2423 VFLNOVGSTLOEBLE-HKKGEMQLEFG-----DLHVDACK--LSEGMQOEN 2467
Db 2680 EYLOSYSKLENELELTGKDKMSFEYKYNKMTAKETELQREHMAOKTAELOEELSGEK 2739
Qy 2468 RRIASTIOLLTKRLKAVVOSKQREITYVLNQEAKLOEKKONKELRYMHNRPSASV 2527
Db 2740 NKLAGEIOLLLEIKS---SKQO-----LKEVLENSLKK-----SLDC 2776
Qy 2528 MEENARLGLILKTVODESKLOSRIKMLNELNLVKDDAMKGEK--VALIQLKLSRN 2585
Db 2777 MHKDOVEKEG---KYREBIAEYQLR-----LHEAKKQIALLD--TKQ 2816
Qy 2586 AEAEINAMOVKLTQKDNLOAMKEIENLOKVNAGAVPYKEIDN-----LTKVYK 2638
Db 2817 YEVEIOTYREKLTSEKELSSOKLEIDLKSS-----KELNNSLKATQIOLEELK 2868
Qy 2639 IEMEKIKY-----SKATQOELIYVLKSC--LEDEKGLRLKLELRAADNDITYCV 2688
Db 2869 TKMDLKLKYNOLKKEKNEAOGKMLLKSSCQOLEEKEKILQKELSQLOAAOQKORT--- 2924
Qy 2689 PRDYOKASTFVYCGSGGIYOSTAMLYOSE---KAAL-----RELSHYKKYHNLST 2741
Db 2925 -----GYWMDTVDELITTEIKELKLEKTEKTEADEYIDLKYSL--- 2963
Qy 2742 MSSSEDRKTRAKSDAHSSHTGSSH---RGSFKTEETRYRGPVTPRSEMPSL----- 2791
Db 2964 LISHEKLEKAKEMLETOYVAHLCSQOSKODSRGSP-----LLGPVVPQSPPIPSYTEKRL 3017

```

```

Qy 2792 -----HIGSPKSESS-----TKRVSPNRSEIYS-----QLYMS----- 2821
Db 3018 SSGQNKASGRKORRSSGIWENGRRPPAPESPSSKSKKAAMVSGIHPAEDTEGETEPEEGL 3077
Qy 2822 -----PGKTGMH-----KHLSPSKVGLH-----KKRALS 2846
Db 3078 PEVYKKGAFDIFGKTSPIYLKRTTMTARTSPRLAOKALSLPLKRENLAESSKPTAG 3137
Qy 2847 PRSEMPTOHVISPGKTG--LHKNLTESTLFDNLSSPCKQKQVQENLNSPKGLFQVSK 2904
Db 3138 GSRSQKVKYKQSPVDSGTLREPTKSVPVNLPERSPTDSPRGLRVKRGRL----- 3191
Qy 2905 SMPYCSQOFUNSKLGDSELTANESNDKSOAENMW 2940
Db 3192 ---VPAKLDWSOLA---VTRVSSALCVSDPW 3219

```

RESULT 2

PCT-US95-16216-1

Sequence 1, Application PC/TUS9516216

GENERAL INFORMATION:

APPLICANT: Yen, Timothy J.

APPLICANT: Ratner, Jerome B.

TITLE OF INVENTION: Nucleic Acid Encoding a Transiently

TITLE OF INVENTION: Expressed kinetochore Protein, and Methods of Use

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dann, Dorfman, Herrell and Skillman

STREET: 1601 Market Street Suite 720

CITY: Philadelphia

STATE: PA

COUNTRY: USA

ZIP: 19103-2307

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/16216

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/353,700

FILING DATE: 09-DEC-1995

ATTORNEY/AGENT INFORMATION:

NAME: Reed, Janet E.

REGISTRATION NUMBER: 36,252

TELECOMMUNICATION INFORMATION:

TELEPHONE: (215) 563-4100

TELEFAX: (215) 563-4044

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 3248 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: not relevant

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ANTI-SENSE: NO

PCT-US95-16216-1

Query Match 7.8%, Score 1148.5; DB 5; Length 3248;
 Best Local Similarity 20.3%; Pred. No. 2.3e-48;
 Matches 715; Conservative 626; Mismatches 1180; Indels 995; Gaps 139;

```

Qy 106 QAIQF-----VRIIOEINREELRVSIMEIYNEVKLLDQDRKKKLEIREDRNVY 161
Db 18 QKIQELGQOLDLKKKEKQOFQDLSLAPQKQTK---VENETGETNLKRENOURLME 74
Qy 162 VADLTEELVMDVEHVIQMIKGEKNRHYGETKMDHSSRSHTIFPMIYESRDNPTUSE 221

```

Db 75 IESLEKTRKOKISHILO-VKESOVN--FOEGOLNSGKOIEKLEOEL--KCKSELESRSQ 129
 QY 222 NCDGAVWVSHLN-----LVYLAGSERASOTGAEGVRLKEGNINSLFILOV--- 269
 Db 130 QAAQAGADV--LWPCMTPOKIFITPTPLTPSQYSGSKYEDLEKYNKEVERKRLKLEVAL 188
 QY 270 -IKKISDGOAGFIVYRDSKILTRILONS-----LGNKATVYIICITTPVSPFD 315
 Db 189 QAKKASOTLPQATMHRD--IARHOASSSVFMOQEKTPSHLSNSOR-----TPIRD 240
 QY 316 -----ETLSTLOFA-----STAKHVMTPIHNEVLDEBALLKRYKILDLK 357
 Db 241 FSAVSFSGELETPRPSRSTIIGKRDANSFFGNSSPHILDOL--KAONOELRNKINLE 298
 QY 358 KOLENLESSSE-----TKQAAMAKEHTOLIAETIKOLHKEREDRIWHLJTNIVASOES 411
 Db 299 IRLGHEKEMKGVNKFQELQOLEKAKVELIEKKEVYLNKCDLVRTTAQYDQASTYIT 358
 QY 412 QODQAVKR-----KRRVYAPGKTONSLHAGVSPDMLSRLPGNF-----SKK 455
 Db 359 ALEOKIKLITLEDLSGQRONAESARCSLEOKIKEKEKEFOEELSRQORSFOTLDQECIQMK 418
 QY 456 AKFS-DMSFPEIDSVCTEFSDFDALSMDNSGIDAEMNL-----ASKYTHREKISLH 509
 Db 419 ARLTQELQOAKNMHVLAELDKLTSVQOLENNLEEFKOKLCRAEQAFQASQIKENELR 478
 QY 510 QSMIDEGOISDSVOPHSSKENOLOYLPKDSGDMAEKRASFKEITSLQOOLQSKKEEK 569
 Db 479 RSMEMKKNMLKSHSQKAREVCHLEALKNIKOCINOS-----QNFADEMKAKNTSQ 533
 QY 570 KELVOSFELKIALEEOISVYKAKNLEMTNREHSINAEVOTDYVEKVEYKREMSYLDGSG 629
 Db 534 ETMLRDLQEKINQOQNSLTLEKILKLAVA-----DLEKRDQCSODILKRR-----EHN 580
 QY 630 YNASNSDLQDSDVQKRLSSHDECIEHRKMLEQKIVDLEFIEINLK-----KSENDK-- 683
 Db 581 IEQLNDKLSTEKESKALLSA-----LELKKRYEELKEKELTFCGWSKSENEKL 630
 QY 684 -OKSEODFMS-----IOLC-----EALMAEKANALELALMRDNFNIT 723
 Db 631 TOMESEKKNLOSKINHLETCIKTQOIKSHEYNERVRLIEMDRMISVEIRMLHVLDSKS 690
 QY 724 LENETLKREIADLERS-----LKENOETNEFELLEKTOKEHEBAOLIHE 767
 Db 691 VEVEFQKLAHYMELQOKAFESQKOKELENNCLTSQITQVEEDL-----EHLQILDS- 743
 QY 768 IGSILKLVENAEKYNLEEDLETKILKLEQETOL----- 803
 Db 744 -----NEIMDKDRCY--ODLHAVEESLRDLKSKSDASLVYTNEDHORSILAFDQOPAMHHSFA 798
 QY 804 -----ALKRKRADN-----LOKRYRNDLSVSM-----GISEKICEBIFQ 838
 Db 799 NITEGQSMSESEKRELEADQSPKSAIILONRVDSLEFISLESOKOMNSDLOKCEBILVQ 858
 QY 839 LKOSLSDAEAVYTRDAQEKCSFLRSENLELKEKMEDTSNMNYN-----OKKASLFF 888
 Db 859 IKGBIE--ENIMKRAEQHOSVAFETSORISKLOEDTISAHONVVAETISALENKEKEIQL 916
 QY 889 EKOLETKSNYKKE-----ADLOKLEQ-----SAPNEIYNIN 921
 Db 917 NDKVETBOAELOELIKKSNHLEEDLSKLEQILSETLSLEKKEMSSIISLNRTEIBELQEN 976
 QY 922 GLI-----AGKVPRLDSHV-----ELEKKVSEFSQOLE----- 950
 Db 977 GYTLKEINASINOEKAMLIQSESEFANTIDERSKISLISLQYKOEKILLILOCEERTGNAY 1036
 QY 951 -----KALEEKNA-----LENEVYTCISEKFLPNEVECKNOISKASEEIMF--LKOGE 998
 Db 1037 EDLSQKTKAOKESKLECLINBCTSLCENR--KNELEQKLEKFAKHOEFLIKLAFAE 1094
 QY 999 HSAIISIKOITIMO-----EQSE-----QIITOLDEVTHTQSVK-----QOTE 1036
 Db 1095 RNONIMLELETVOALRSEMTDNONNSKSEAGIKOIEIMTLKEONKMKREAVNDLLQEN 1154

QY 1037 EOYLEMKRMD--DLFEKTIANKSEADLLREMENTKCTME-----SVE 1078
 Db 1155 QLMVKMYKTRKHCQMLSESPINNYSVKERESEKNOCNFRQMDLEKVEKISLDSYNALVOLE 1214
 QY 1079 VKIADTKHELEETIRDKBOLHKEKKYFFQAMQTFPTPLSDSRLPSPKIVGNSODPIE 1138
 Db 1215 AMLNKELEKLOESBEKECLOHE-----LOTI-----RGDEITSNLDQMSQELISGL 1261
 QY 1139 NDY-----HNILATERNNIMVOCLETFRNSLEQOVIDNTLOLSQASIE- 1185
 Db 1262 KDCEIDAEEKYISGPHLESTSQNDNAHLQCSIQTKMKLNE--LEKICEIILQAEKVEL 1317
 QY 1186 -----KSDLOKPRODEGEVKLLMEMLK----- 1212
 Db 1318 VTELNDRSBECTITRKAABEVGKLLNEVKILINDSDGLHGEVLVEDIPGEBFGEQONEH 1377
 QY 1213 -----HLTDSOLSIKLOLENLEVTEKLOTLQOEMKNITIRNELOTNFD 1258
 Db 1378 FVSLAPIDESNSYBHLT--LSDKEVQMHFAELQEKFLSLQSEKILHDOHQMSKME 1434
 QY 1259 LKAHDSILKODLSINIQSISTODELPAQOELREQOQYVDSFRQOLLDCSVGISSEFND 1318
 Db 1435 LOYVDSILK--AENLVLS--TNLNFQDLYKEMQL--GLEEGIVPSLSSCVPDSS 1485
 QY 1319 AVANOEKVSIGEVNSLOSEMIRGERDELQTSKALVSELE-----LLRA 1362
 Db 1486 SLSS-----SLGD--SSYFRALLE-----QTDGMSLSLNLGAVANANOCSDVEVFCSSLQ 1533
 QY 1363 HVSVEGENLEITKINGLEKEILKSE--ESEVYKSMLENEKEDNNKLR----- 1410
 Db 1534 YVDSLKAKENLVLTNLNFPQDLYKEMQLGELGIVPSSSCVPOSSSLSGDSFYR 1593
 QY 1411 --EQAEEYS-----SKENQFSLEEV-----GSSQKLYBEI--E 1440
 Db 1594 ALBQOTDMLSLSLBSVAVANOCSDVEVFCSSLOEBENLTRKTPSPAPAGVELLSICE 1653
 QY 1441 VLAQILAAEBERLE--IKRDYFELVQANTNLVE--GKLETPLOADEH--EDSIDRRS 1493
 Db 1654 VYROSLEKLEKESKQIMKKEKIELOELQELSSERQDLRLROYLSENQOQOKLTSV 1713
 QY 1494 EMEIKVYLGKLEKRNQYLE--RQOEKLELSNKLEILOKEMETSVYLLKDLQOKLES 1549
 Db 1714 LEMESKILAAKKOTELOLSLEVARLOLOGLDLSR-----SLGIDTEBAIOG 1762
 QY 1550 LLSENIILKENIDITLKHSDTOALOK-----TQOETOL-----AKNULATAAGD 1594
 Db 1763 -----RNSCDSIKENHSETTERTPKHDYHOCIKQKAOQDNLNLEKITEGAVKPTG 1815
 QY 1595 NCPITQOKET-----SADCVHPL-----EKKILLITTEELHOK 1626
 Db 1816 ECGSEQSPDITNBPPEBDKTQGSSEICISLFSGPNALVPMDFLQNOEDIHNLQLVKFT 1875
 QY 1627 TNEOEKLEKKNLEBOAQVELKCEVEHLKMSIESASSLESLOHEKHPDEOLLA----- 1681
 Db 1876 SBNENLRILHVEIBDR-----KVESLNLNEMKELOSKL-----HLOEVDLMKTEAC 1921
 QY 1682 --LKQOMVYVTOEKKELQOTHEHLTAEVDLKENIEL--GLNFKNBAQ--OKTTKEOCLIN 1736
 Db 1922 TELEKIVGELKKNKNSDLSKLEVFSCDQOELLQRYVETSGIANSDLEMHADKSRREDIGN 1981
 QY 1737 ENK-----ELBOSORLOCE-----IEELMKSLKDESALETLKSEQRY 1776
 Db 1982 VAKVNDSMKREFLVDENELSRIRSEKASIEHBALVLEADLEVQOETKLEKONENKQV 2041
 QY 1777 I--NLNOEMEMVMELEKNSORTVYABR--DOLODDRESVENSIER----- 1821
 Db 2042 IYCLEBELSVYSEKQOLGELDITMSKKTALDQOSEKKKEKTO--ELBHSQECILACIOV 2100
 QY 1822 -ODDLKRAQALQOQKQVQELTSQISVLOEKITSLLENOMLYNAVATVKETLSERDNLNS 1880
 Db 2101 AEAVEVKEKTELLQTLSSDVSEILKDKTHQEKLOLSLEKXSQALSILKCELENOQALNKE 2160

Db 163 HSFANIGEGSMPSESECRLEADOSPKNASAILQNRVDSLEFSLESOQOMN-----SD 216
 QY 756 TOEHEAQLIHEIGSLKLYENAMNTQNIIEBLETYTKILKEOIOALERRADN--- 812
 Db 217 LOKOCE--ELVOIGKEIEENIMAKAOMHOSFV--AETSQIRISKLOEDTSQAONVAVETLSA 273
 QY 813 LOKKVRNFDLSVSGDSEKICEEFIFOLKOSLSAENAVTRAOKACSTLRSENELEKME 872
 Db 274 LENKEKLOLLNDKVEEIO--AETIOELKS---NHLEDSLEKLOLL--SETISLEKKE 326
 QY 873 DTSNMVOKERKASLFEKOLETERKSNNYKKEADLOKE---LOSAPENIYVINGLAKV 928
 Db 327 SSTISLNRKRE-----IEELTQENGTLKEINASLQNEKMLIOKSEFANYID----- 373
 QY 929 PROILSVELEKVKSESKOLE-----KALEKKA-----LE 960
 Db 374 -----EREKSISETSDYKOEKILLORCEETGNAYEDLSOKYAAOEKSKLEBCL 425
 QY 961 NEVYCLSEYKFLPNEVECLKNQISKASEEIML--LKQGEHSASTISKOETIMQ----- 1012
 Db 426 NECTSLQENR--KNELQLEAPAKHEQEFITKIAFAEERNQNMILEYVQOALRESEMT 483
 QY 1013 -----EOSE-----QILOLDEVYHTQSV---QOTEBOYLEMKRMD--DLEKYIR 1055
 Db 484 DNQNSKSEAGGLKOEITMTLKEEQNKQKEVNDILOENBOLMKYMKTKHCQNLSEPIR 543
 QY 1056 NKSABEILLEMENLKQTE-----SVEKXADTKHELEETIRDKCOL 1099
 Db 544 NSYKERSERQOCFKPOMOLEYKELISLSTYNAOLVLEAMLRKELKLOESEKELQ 603
 QY 1100 HEKKYFQOAMQETPIPIPLSDSPRSKLVGNSQDPIEINDY-----HNILAL 1147
 Db 604 HE-----LOTI-----HGDETSNLOMQOGEISGLKDCETDAEKYISGHETLS 650
 QY 1148 ATERNNIMVCLTEBERNSLKEVDYDNLQLOSLAOSTE-----KSDLOKPKODLEEG 1199
 Db 651 QNDANHALQCSLOTMTNKLNE-----LEKICETLQAEKELYELNDSRSECTITATRKAAEE 706
 QY 1200 EYKLLLEMELLK-----LEKICETLQAEKELYELNDSRSECTITATRKAAEE 706
 Db 707 VGKILNEVKILINDSGILHGEIVEDIGGEGEOPNEOHVSLAPLDESNEYELT---L 763
 QY 1220 SIEKLOLENEVEYKLOTLQOEMKNITIERNELOTFNEDEKKAHDSLKODLSENIEOSIE 1279
 Db 764 SDRKVOHMFALQEKFLSLQSEHKILHDOHQMSKMSSELQTYVDSLK---AENVLVS-- 818
 QY 1280 TOEELRAOEBLROKOLVDSFRQOILLDCSVGISSPNHDAVANOEKVSLGEVNSLOSEM 1339
 Db 819 --TNLRNFQDGLVKEML--GLEGLVPSLSSSCVPDSSSLSS---SLGD--SSFYRAL 868
 QY 1340 RGEDELOQSCALVSELE--LTRAHVSEGENLEITRKILNGLEKELIGKSEES----- 1392
 Db 869 E-----QIGMISLISLMEGVANSQSYD---EYF--CSSLOEMLTRETPTSAPAKG 916
 QY 1393 -EYKLSMLENLKEDNNKLIKQAEYESSKENOFSLEEVSSQKLVDEIEYKALQAAEE 1451
 Db 917 VEELIESLCEYVROSLKELEKMEGQIMKNK-----BIOLEQLLSERQ 961
 QY 1452 RLEIKNDYELVQNTANTNLVEGKLETPLOADHEDSIDRSEMEKLYGKEKERNQYL 1511
 Db 962 ELDCLRQOYLS-----ENEOQOKLTVTYLEMESKLAEEKQTOELS 1003
 QY 1512 LE-----RLOEKELYSKLELLOKEMETSVLLKDDLOOKLESLSSENIILKENIDTTLKH 1567
 Db 1004 LELEVALLOLOGLDLSR-----SLGIDTDEDALOG-----RSCSDISKHE 1045
 QY 1568 HSDTOALQK-----TOEELQ-----AKNLAIASDNCPIQOERET----- 1604
 Db 1046 TSETTEPRHDVQICODKAOQDNLNDIKETGTALKPREGSGSGSPDTNPERPGED 1105
 QY 1605 ----SADCVHPL-----EKKILLTEELKOKTNEOKLLHKKNELBOAQ 1644
 Db 1106 KTQSSSEICISLSPSGNALVPMDFLQNOEDIHMLQLRVKEYSNENRLLTHVEDRDR-- 1163

QY 1645 VELKCEVHEMLKMEISKSSLESLOHEKKHDTBOQLA-----LKQOQVYTOEKKELQ 1697
 Db 1164 ----EYESTLNMKEKELDKL-----HLOEQVLMKTBACLETETKLYGELKKESSDS 1211
 QY 1698 QTEHNTAEVDHLKENIEL--GLNFRNEAQ--QKTEBOCLNENK-----BLEOS 1744
 Db 1212 EKELEYSCDHQELLQVETSEGLNSDLEMHADSSNEDISQDNYAKVNDMSKKEFLVEYN 1271
 QY 1745 QHRLQCE-----IEFLMSLKDKESALETLKESEQKY--NINQEMENYLMMEEL 1793
 Db 1272 LSRIRSEKASIEHEAVUYLEADLEVYQTEKLEKDNENKQVYVCEBELSVTSEBNOL 1331
 QY 1794 KNSQRYVIAER---DOLQDQRESVEMSTFT-----QDDLRKAOALQOQKDK 1838
 Db 1332 RGEIDTMSKKTAYADDOSEKMEKETO--ELFSHOSSECHQVQVAAVEKTELLQTLSSD 1390
 QY 1839 VOELTQISVLOERKISLENNQMLYNVAYKELTSEBDLQNSQKHLFSETITLSLKEK 1898
 Db 1391 VSELLDKTGHQKLOSLKESQALSLTKCELENQLOLNEKRELLVKESESQARLSIES 1450
 QY 1899 EF-----ALBOAKDKAARKTIDITEKISNI-----BEOLLQOATNL 1937
 Db 1451 DYERLWYSKALAEALVEKGFALRLSSTQEBVHQLRGIEKLEVRLEADOKLOLHAERK 1510
 QY 1938 KETLYERESLQCKEOLATHEHLRETILKSDALGKMEQERDEANKVYALTEKMSLSE 1997
 Db 1511 KERERENDSL--KDKVYENLERELOMSEENQELVY--LOAENSKA-----EYETLK 1556
 QY 1998 EQJNEVNTYKKEG-----GEKETFYLOPRSKQOSSOMELRSLTKYDLOLEBAEK 2050
 Db 1557 TOJIEEMARSLKVELDLVTLRSEKENLTKOIOEKOGOLSELIDLKSFSLKEKEKQAE 1616
 QY 2051 EISEAT-----NEKNTLAKISSL--BEELIQ--NASILNAYSERENLRHSKOIYS 2099
 Db 1617 QIKESKTAVEMLONOJKELEBAVALCGOETIMKATQESLDPPLEBEHOLRNSIEKTRA 1676
 QY 2100 ELE-----QSLTLKSRHFAQSKREKDEAVNKIASIAEIKILTKEMD---EPRDSK 2150
 Db 1677 RLEADEKQKQVYLOQLK-----ESEHNDLKKQVELELELEIARNOHMALEBENSK 1731
 QY 2151 ----ESLOEGSSHUSEELCTYKTELQOMLKQOKEEDINNKLAKYAVEDELLQHLSSLEKOLD 2207
 Db 1732 GEVEYELTKAKIEGTQOSLGRGELDVVYTRSEKENIYTELVEOERISL----- 1779
 QY 2208 QIQMELNEKRLNRYELCEKADIMEKEISVLRIMQNEPOQEREDVAAERMDLLESRNOETOE 2267
 Db 1780 ----EITNSSFEN-----ILOEKBOEVOMKESSTRAMENLOT---QIKE 1817
 QY 2268 IMEKISAVYSEOHTLSSLSSELQKETEAKHQMNLKESLSSLSRSFGSLQTEHYKLN 2327
 Db 1818 LNERVALHNDQ-----EACK-----AKEONLSSQVCELELEKAOQL 1854
 QY 2328 TOLOTLNKKRYVYRPAAYKEDSLTKDYKEDLAADOKRDE---RLQLOQCELOGRK 2384
 Db 1855 QGDKAANNYTVLOSSY---NGLIQVEDGQOKLEKQESLSRKLNQLODOEOLVSKL 1909
 QY 2385 SDSASELEKE---CELEFLNELFLKFRANITQSVODFSEVQVFLNOVGSTLOEBELE--HK 2439
 Db 1910 SQYBGEHQLMKEONLEIRNLITVELEQKIQVLOSKNASLODTLEVLQSSYKNLENELETK 1969
 QY 2440 KGFQWMLSEEG-----DLHADYAK---LSEGOQENRRIASTIQLTKRLKAVY 2485
 Db 1970 MDKMSFVEKYNKMTAKETELQREHNEHAQTAELQELSEKRNRLAGEQLLLEIKS-- 2027
 QY 2486 QSKIOREITVYVLOFEAKLOEKKQONKELRRMHNHSPASVABEENAKRLGLITKYODE 2545
 Db 2028 -SKDQ-----LKEITLENSELEK-----SIDCHNKQOKEG---KYREE 2063
 QY 2546 SKTKQSRKMLENLEMLYKRDAMHKGK---VALILOKLSRNAEALANOVULTKRODN 2603
 Db 2064 IAEYQLR-----LHEAKKHOALLLD--TKQYEVETQYTERKLTSEKEC 2106

QY 2604 LOAAKEIENLOKMAVAVPYKEEIDN-----LTKVYKIMEKIKY-----SK 2648
 Db 2107 LSSOKLEIDLKSS-----KEELNNSLKATQIILEELKTKMDNKIYNOLKKEER 2158
 QY 2649 ATDOELAVLKSC--LEDEBGLRLKEBELRRQAQDNDTVCVPKDYOKASTFPVTCGGGS 2706
 Db 2159 AOGKMLIKSCQLEBEKEILOKELSOLAQAEOKT-----2196
 QY 2707 GIYSTAMLVQSE-----KALE---RELSHYKKKHYHLSMTSSSEDKRTAKSDAHS 2759
 Db 2197 GTVADTVDELTYTEIKELTEKEADEYLDKCSL---LISHHELEAKIMLETOV 2253
 QY 2760 SHTSSH-----RGSFKHETERYRGVTPPERSEMPSL---HLGSPKSESSTKRVVS--- 2808
 Db 2254 AHLCSQSKODSKRSP-----LLGPVYPGSPPIPSVTEKRLSSGQNKASGRORSSGIW 2307
 QY 2809 -----PNNSIYS-----OLVMS-----PGKTM 2827
 Db 2308 ENGSGPPAPPESEFSKSKKAVMGIHPAEDTEGEFEPEGLPEVYKKGFADITGKTSP 2367
 QY 2828 H-----KHILSPKVGHL-----KKRALSPKREMPTOHVISPKRTG 2864
 Db 2368 YILRTTMAFTSPRLAQAQALSLISLGENIAESSKPTAGSGRSQKVYVAORSVPDSG 2427
 QY 2865 --LHKNLTESTLFDNLSSPCQKQKVQENLNSPKGLF-----DYKSKMPCPSQ 2912
 Db 2428 TILREPTKSPVNNLPERSPDTSPREGIYKRGRLVPSPKAGLESKSENCQVQ 2482

RESULT 4
 US-09-572-191-2
 ; Sequence 2, Application US/09572191
 ; Patent No. 6355466
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; APPLICANT: Sakowicz, Roman
 ; APPLICANT: Wood, Kenneth
 ; TITLE OF INVENTION: No. 6355466 motor proteins and methods for
 ; FILE REFERENCE: 1017
 ; CURRENT APPLICATION NUMBER: US/09/572.191
 ; CURRENT FILING DATE: 2000-05-17
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: FaastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 1388
 ; TYPE: PRT
 ; ORGANISM: Human
 ; US-09-572-191-2

Query Match 7.13; Score 1051; DB 4; Length 1388;
 Best Local Similarity 24.58; Pred. No. 5.2e-44;
 Matches 417; Conservative 304; Mismatches 548; Indels 434; Gaps 60;

QY 2 SEGDAVAVCVRVRLIORE---OGDQANLQWKAGNNTISQVDT---KSFNEDVFNSSHE 55
 Db 22 NEGDAIVFYRIRPARRSAGSAGEQ-NLCISVLSSTSLRHSNPEKRTFFDHVADVT 80
 QY 56 STSOIOELVPIIRSLQGYNGTIFAYGOTSGKTYTMG-----TPNSLGIIOAI 108
 Db 81 TOESVFATKSIYESCMGNGTIFAYGOTSGKTYTMGSESDNFSHLRGVIRSF 140
 QY 109 QEVYKIOEL-----PRRELLRYSYMEIYENYKDLCDRRKKPLEIRDFKRNYYA 163
 Db 141 EYFSLIDREKRAKAGSKSLCKCSFLEIYNEQIYDLL--DSASAGLYLRHRIKGVFV 198
 QY 164 DLTEELVAVEHVLIQIKKEKKNHYGETKMNHSSRSHTFRMIVESRD-NDPTSEN 222
 Db 199 GAVGOVYTSAAEAQVYLSGGGRNRRAVASTSMNRESSSHAVFTTISMESKEINIV--- 255
 QY 223 CDGAVVYSHNLVDLASSERASQTAGVRLKEGCNINRSLIFIGVYIKLSDGQAG-G 280
 Db 256 ----IRTSLLNLVDLASSERQKTHAGMRLEKAGNINRSLISLGGVYITALVDVGNCKR 311

QY 281 FINRDSKLTIRLONSLGNAKTVICTIPVS--FDETLSTLOFASTAKHVRTPHYNE 338
 Db 312 HCTYRSKLTFLVLRDLSLGGNAKTAIIANVHPGRCFGETLSTLNFQRAKLIKNAVYNE 371
 QY 339 VLDEALRYRKETLIDKOLENLESSSETKAQAMAEETTOLAEIKOLHKEEDRIW 398
 Db 372 --DVGNSVQLOAEYKRLKEQALASGQTPPESTLFTDKKTYMVEFQ-----EAMLF 424
 QY 399 HLNIIVVASSQESQODQYKRRRYTMAPKIQNSLASVSDPMLSRLPGNSKRAK 458
 Db 425 F-----KKSQEKK-----SLIEKYQLDELDLYLKEKF 452
 QY 459 SDMSFPEIDSVCTEFDFDALSMDNSGIDAEWNLSKVTYHREKTSLSHQMIDEGOI 518
 Db 453 -----IQSNMIVKFR-EDDIIRLEK--LH-----474
 QY 519 SDSVQFHDSSKENOLQYLPKDSGMAECRKASFKEITSLQOOLSKKEEKELVQSPFL 578
 Db 475 -----KESRGFLPEEQDRLL-----SELNREIOTLREQLEHHRVAKYAMENSL 520
 QY 579 KIALEBOLSVYKANLEKVTNSRHSINAQVQDVEKEVYKREMSVGLDSGYMASNSDLQ 638
 Db 521 R-----EENRRRLRL--LEPVKRAQE-----538
 QY 639 DSVQGRKLSSHDCEIHRKMLBOKIYDLEFIENLUNKSENDKOKSSBODFMESTQLC 698
 Db 539 -----MDQTTAKLEKASEISEKMSKSD--NOOGFSPKRAKE 574
 QY 699 EAINAEKANALELALMDFNFIENETLKREIADLSERLKEN-OETNEFEILEKETQ 757
 Db 575 PCLFA-----NTEKLAQLOQIOTLNNSKQYEEF-----KELT 609
 QY 758 KEHEQOLIHEIGSLKYLEVNAEMTQNLIEDLETYKILKQETOLAEIR-RADNIQ-- 814
 Db 610 RKQOLELESELOSLOKA-----NLNLENLEA-TKACKQAEV--SQLNKIHAEYTKII 659
 QY 815 --KVRNF-----DLSVMSGDESKCEIIFQKOSLSDAEVATRDQKCSFLRSNLT 865
 Db 660 TTPTKAYOLHRRPVKPSPEKSGSL-----YTQNSSTLNDIINEPVPE-----706
 QY 866 ELKEMEDTSWYNOKEKAASLFEKQLETERSNYKMEADLOKE-----LQSAFNEINY 919
 Db 707 -----MNEQAFEAISEBELTVOEQSALQKDEEBHKLKQOHVHDKLH 752
 QY 920 LNLGAGKVPRLDSRYELEKVVSEFSQOLEKALEKNALENY-TCLSEYKFLPNEYEC 978
 Db 753 HSTQM-----OELFSSERI-----DWTQOEBELLSQLVLEKQOLEOTKNDPFLKSEVHD 802
 QY 979 LKNOISKASEEIMLKQF-----GEHSASITISKOELIMOESQOLQOLDDEVTHQSK 1031
 Db 803 LRVVLHSHQKELSSVCLTEYSSFTKNOEKEFKLSERHMHVOLDNLKLENEKLUESKAC 862
 QY 1032 VOQTEEOYLEMKKMHNDLFEFYIRN-KSEADLLREMENTLGTMESEYVK-----IA 1082
 Db 863 LQSDYDNLQELMKKEIDQLSRNLQNFKENETLSDNLNMLLEAEKERNNKSLQPEE 922
 QY 1083 DTKHELEETIDKQOLHEK-KYFFQAMQITFTPTLSDSLPSPKLYEGNSQDPIELNDY 1141
 Db 923 DKENSSEIKLVLEAVROKQETAKCEQOMAKQKKEESL-----963
 QY 1142 HNLALATERNNINWCVCTEENSILKEOYIDLPQLOASQISKD-LQPKQODEGE 1200
 Db 964 -----LATER--VSSLEKSDSDKKVADAMNOIOELRSSVCEKTEFTIDTLKQELDIN 1016
 QY 1201 VK--LLELLEMLKGHLTDSQLSIEKLOLENLEYTEKLOTQOEBMKNTITERNLQTNFE 1257
 Db 1017 CKYNSALVD-----ESVVLKKQEVOLDLKEFLR-----LRLISE 1054
 QY 1258 DLKAEHDSLKODLSENIBO-----STFODELRAAOEELREOKOLVDSFROOLDOSV 1310
 Db 1055 DI--ERDMCLCIDLHATEOLNMLTEASKKHSGLLOSQOEELTKREALIOELHRLT-----1107

QY 1311 GISSPNHDAVANOEKVSLGEVNSLOSSEMNGERDELQTSKALVSELELLRAHVYSVEGE 1370
 Db 1108 -----NOKK-----EVEEQKKNEYNPKMQL-----EHMDSAAE 1137
 QY 1371 N-----LEITKKLNLEKEILKSESEVLSMLENKEDNNKLE--QAEYS 1417
 Db 1138 DPOSKPPPHQTHAKLETQEOEI-----EDGRASSTSEHLVTKLNREVRVNAEILR 1193
 QY 1418 SKENQFSLEEVFSGSOKLVEIEVILKAOLKAERLEIKDYPFELVOTAN-----1468
 Db 1194 MKEQLOREMENLRLESQOLIEKNMILLOQOLDIKKQKENSQDNHNDQOLNKEQESITKE 1253
 QY 1469 ---TNVYEGKLETPLOADHEDSDIDRSEMEIVLGEKLEKRNQYLLERLOEKEILSNK 1525
 Db 1254 LAKSRIVEEMIK--MKADLEEVOSALYNNKEMECILRMTEDEVERTQLESKAFQEREQLRSK 1311
 QY 1526 LEIQKEMETSVLLKDLQOKLESLSSEN--IILKENIDITLKHSPTQOALQTOEQLO 1583
 Db 1312 LEMYEERERISQEMELRKQVECLAENGLVGHQNL-----HOKIQYVVRLLKENVR 1365
 QY 1584 LAKNLAIAASDNCPTIOEKETSA 1606
 Db 1366 LAETETKLRAEVNFLEKERRSES 1388

RESULT 5
 US-09-723-262-2
 ; Sequence 2, Application US/09723262
 ; Patent No. 6379912
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; APPLICANT: Sakowicz, Roman
 ; APPLICANT: Wood, Kenneth
 ; TITLE OF INVENTION: No. 6379912el motor proteins and methods for
 ; FILE REFERENCE: 1017
 ; CURRENT APPLICATION NUMBER: US/09/723,262
 ; CURRENT FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: US 09/572,191
 ; PRIOR FILING DATE: 2000-05-17
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 1388
 ; TYPE: PRF
 ; ORGANISM: Human
 US-09-723-262-2

Query Match 7.1%; Score 1051; DB 4; Length 1388;
 Best Local Similarity 24.5%; Pred. No. 5,2e-44;
 Matches 417; Conservative 304; Mismatches 548; Indels 434; Gaps 60;

QY 2 SEGDAAVGVVRVPLQRE---QGDQANIQKAGANTTQVDGT---KSFNDRVFNSSHE 55
 Db 22 NBDADAKVVRIRIPRAERGSADGEQ-NLCLSVLSLRLHSNBPCKFTFDHADVDVT 80
 QY 56 STSOIYQEIADVPIRSALOGYNGTIFAYGOTSSGKTYTMAG-----TPNSLGIIPAI 108
 Db 81 PQESVAVATVAKSIVESCNGYNGTIFAYGOTSGKTYTMAGSSEDNSHNRGVIPNSF 140
 QY 109 QEVFKITQET-----PNRELLRVSYMEIYNETVDLDCDDRRKRPLEIREDFNNVYA 163
 Db 141 EYLFSLIDREKKEKAGAKSFLCKSPFIETIYNEQIYDLT--DSASGLVLRHEIKKGVYV 198
 QY 164 DLTEELVMPBEHVIQMIKKGEKNRHYGETKMDHSSRSSTITIRMAVESDR--NDPTNSEN 222
 Db 199 GAVEQVYVSAAYOVLSGGWRNRVASTSMNRESSRSHAVFTITIESEKSNETYN---255
 QY 223 CDGAVVASHLNLVDLAGSERASQTAGEVYRLKEGCNINRSFLITLQGVYIKKJSDQAG--G 280
 Db 256 ---IRTSILNLYDLAGSERQKDTAEGMRLEKAGNINRSLSCLQGVYITLAVDNGKQR 311
 QY 281 FLYNDSKILTRLQNSLGSNAKTVIICITTPVS--FDETLSTLOPASTAKHVRNTPHYNE 338

Db 312 HVCTYRDSKTLFLRLDSISGNAKTAITIANVHPSRCGETSLTNFQRAKLLINKAVNE 371
 QY 339 VIDEDEALLKRYREIIDLKQOLENLESSSETKQAAMAKEHTOLLAIEIQLKHEREDRYM 398
 Db 372 --TQGVNSQLOAEVRYKRLQELASGCPPEPSFLTRDKKKNVMEYEQ-----EAMLF 424
 QY 399 HLNIVYASQESQDOQRYKRRRYTWAPKQIONSILHAGVSDPMLSLPGNFSKAKAF 458
 Db 425 F-----KKSQERK-----SLIEVQTLELTLTKKEF 452
 QY 459 SDMPSEPIDDSVCTEFSDFDALSMDSNGIDAENWLASKYTHREKTSIHQSMIDFGI 518
 Db 453 -----IQNKMIYKVR--EDQILRLEK--LH-----474
 QY 519 SDVQVQHDSSKENQLOYLPRKSGDMAECKRASFKEITSLQOOLQSEKREKELVQFEL 578
 Db 475 -----KESRGFLEBEDRLL-----SELNRIQTLRQDIEHHPVAKYAMENHSL 520
 QY 579 KIALBEQSLVYKAKNLEMTYNTNSREHINAQVQDVEKVEYVRKEMSVLDSGYNAANSQDQ 638
 Db 521 R-----EENRRLR--LEPYKRAQ-----538
 QY 639 DSSVDGKRLSSSHDECTEHRKMLEQKYDLEEFTEENLKKSENDOKSSEODPMESIQC 698
 Db 539 -----MDAQTIAKLEKAFSEISGMEKSDK---NOGFSPKAQE 574
 QY 699 EAIMAKKANALELALMRDNFNILNETLKRLEIADLERSLKEN--QETNEFEILEKETQ 757
 Db 575 PCLFA-----NTEKLAQLOLQOTELNNSKOEVEE---KELT 609
 QY 758 KEHEQOLHEIGLSKLKYVENAEMYNQNEEDETETKLLKEQEIQALERK--RADNLO--814
 Db 610 RKRQLESELSQLOKA-----NLNLENLLEA--TKACKREY--SOLNKHATLTKIT 659
 QY 815 ---KAYRN-----DLSVMSDSEKLCETIFQLKQSLSDAEATYRQAKCECSFLRSEN 865
 Db 660 TTPYTAQVLOHSPVPKRLSEMGFSGL-----YTQNSSIJDNDILNEVPPE-----706
 QY 866 ELKEKEDTSNMYNOKERAKSLFEKOLETEKSNYKMEADLOKE-----IOSAFNEINY 919
 Db 707 -----MNDQAFEISELRTVQPOMSALQAKLDEEHKMLKQOYHVKLEH 752
 QY 920 INGLIAGVPRDLRSVLEKRVSEFSKOLEKALEEKNALENY--TCLSEYKFLPNEVEC 978
 Db 753 HSTOM-----OELFESSERI-----DMTKQOEBELSQLNVLEKQOETQKNDPLKSEVHD 802
 QY 979 LKNOISKASEETMLKOE-----GEHSATISQOEIIMQOSEQIQLTDEVTHTQSO 1031
 Db 803 LKRVYLSHSDKELSVKLEYSSEKTIQOEKEFNKLSRHHVQIQLDNLLENEKELLESKAC 862
 QY 1032 VQOTEQOYLEKKMHDDLEFYKIRN--KSEADLLREMENTLGTMESYEVK-----IA 1082
 Db 863 LODSDYDNQOEMKFEIDQLSRNLQWFKKENETLMSDNLNMLLELAEKERNKKSLOPEE 922
 QY 1083 DTKHELEETINDKQDLHEK--KTFPQAMQITFPTPLSDSIPPSKLYEAGNSODPIEINDY 1141
 Db 923 DKENSKEIILVLAVRQEKQETAKCEQOMAKYQKLEESL-----963
 QY 1142 HNLLATERNNNIMWCLETENSLKEQVYIDLNTQLOSLOASIKSD--LQPKQDLEGE 1200
 Db 964 -----LATEK--VYSSLEKSDSDKKVYADLMMQIOELRSSVCCKETITIDTLKQELNDIN 1016
 QY 1201 VK-----LLEMLLKGHLTDSQLSIEKLQLENTLEVTEKILQTLQEBMKNTITERNELQTNFE 1257
 Db 1017 CKYNSALVDR-----ESRVLIKQOEVITLDLKEFLR-----LRLISE 1054
 QY 1258 DLKAEHDSIKODSENIEQ-----SIEQDELRAQAOEELREKQVYDSFROQLDCSV 1310
 Db 1055 DI--ERMDLCEDLAHATEQLNMTLEASKRHSGLQSAQEBELTKREALIQLQHLT-----1107
 QY 1311 GISSPNHDAVANOEKVSLGEVNSLOSSEMNGERDELQTSKALVSELELLRAHVYSVEGE 1370

Db 1108 -----NOKK-----EEVEOKKNEYNFKMKQL-----EHYMSAAE 1137
QY 1371 N-----LEITKRLNGLEKEILGKSESEVILKSMLENKEDNNKKE--QAEYS 1417
Db 1138 DPOSPKPPHQTUHLAKLETQEOEI-----EDGRASKTSLEHLVTKINEDEVKNAELR 1193
QY 1418 SKENOFLEEVFSSOKLVEIEVLKAKOAKAEERLEIKRDYDELVOTAN-----1468
Db 1194 MKQOLREMEMLRLESQOLIKEMNLLOGLDDIKKOKENSQNHDPNOOLKNEOESIKER 1253
QY 1469 ---TNLVEGLKLETPLOADHEDSIDRSEMEIKVLGEKTERNOYLTERLOEKELETSNK 1525
Db 1254 LAKSKIYEMLK--MKADLEEVQSALYNKMEECIRMTDEVEROTLSKAFQEKELRSK 1311
QY 1526 LEILOKEMETSVLLKDDLOOKLESLSN--IILKENIDTTLKHSPTQAOLOQTQOELQ 1583
Db 1312 LEEHYERERTSOEMEMLRQOVCEIAEENGLVGHQNL-----HOKIQYVLRKKNVR 1365
QY 1584 LAKNLAASDNCPTQEKETSA 1606
Db 1366 LAETETKRAENVFLKKEKRSSES 1388
RESULT 6
US-09-723-219-2
Sequence 2, Application US/09723219
Patent No. 6391613
GENERAL INFORMATION:
APPLICANT: Beraud, Christophe
APPLICANT: Sakowicz, Roman
APPLICANT: Wood, Kenneth
TITLE OF INVENTION: thelr use
FILE REFERENCE: 1017
CURRENT APPLICATION NUMBER: US/09/723, 219
CURRENT FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: US 09/572,191
PRIOR FILING DATE: 2000-05-17
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 2
LENGTH: 1388
TYPE: PRF
ORGANISM: Human
US-09-723-219-2
Query Match 7.1%; Score 1051; DB 4; Length 1388;
Best Local Similarity 24.5%; Pred. No. 5, 2e-44;
Matches 417; Conservative 304; Mismatches 548; Indels 434; Gaps 60;
QY 2 SEBDAYKVCYRVPRPILORE---OGDDANLQWKAGNNTISQVDT---KSFNEDRVFSHE 55
Db 22 NEDDAIKVFRIRPRRPSRSGSADGEQ-NICLSVSLSTSLHSNPEKPTTFDVAADVDT 80
QY 56 STSQIOEIVLPIRSLAQNGTIFAYGQTSQKTYTMAG-----TPNSLGIIPQAI 108
Db 81 TQSSVATVAKSTVESQMSYNGTIFAYGQTSQKTYTMAGPSPDSFNNKRVIPRSF 140
QY 109 QEVFKITQEI-----PNREFLRVSYMEIYNEVTKDLCDDRKKKPLEIFEDFNRYVA 163
Db 141 EYFLSLIDREKERAGAKSFLCKSCFEIYNEQIYDILL--DSASAGLYIRHEIKKGYFVV 198
QY 164 DLTEELMVPEHYIOWIKKEKNRHYGETKMNDSHSHIIFRATVSROR--NDPTSEN 222
Db 199 GAVEQVYVTSAAEYVYLGGWRNNRRAVASTSMNESHSHAVFTTTESKSENIYV--- 255
QY 223 CDGAVWVSHNLVYLAGSEASOTGAGVRLKGCNINRSLFTIGYIKKLSQOAG--G 280
Db 256 ----IRSLNLVYLAGSEQOKDTHAGMRLKNGINRSLSCIGQVITYLAVYGNKOR 311
QY 281 FNYRDSKLTIRIIONSLGNKATVYICTTPVS--FDELSTLOFASSTAKHVRTPHNE 338
Db 312 HVCYRDSKLTFLRLDRLSIGNAKTAIITANVHPGSRCEGTSTLNFQRAKLTKKAVYNE 371

QY 339 VLDEBALIKRYRKREILDLKCOLENLESSSETKAOAMAKEERTOLLAEIKOLHKREDRIW 398
Db 372 --DTQGVNSQLOAEVKRLKEQLAEIASQGTPESEFLTRKQKTYMEYFQ-----EAMLF 424
QY 399 HLINIVVASSQESQODORVKRRRVYAPKGIQSLHNASGVSDPMLSRLLGNFSKAKF 456
Db 425 F-----KKSQDEKK-----SLIEKYTQLEDLTLLKKEF 452
QY 459 SDMPSPFEIDSVCTEFSDFDALSMDNSGIDAEMNLASKVYTRERKTSLSHQMIDRGOI 518
Db 453 -----IQSNKIYKFR-EDQITLLEK--LH-----474
QY 519 SDSVQFHDSSKENOLOLYIPKDSGMAECRKASFEKEITSLQOOLOQSEEEKELVQSEFEL 578
Db 475 -----KESRGGFLPEEODRILL--SELNREIOTLREQIEHHPRAVKYMEHNSL 520
QY 579 KIALEBOLSVYKAKNLEKVTNSRHSINAEVQTVDEKEVYVAKEMSVLDSGYNANSNDLO 638
Db 521 R-----EENRRLRL--LEPVKAQE-----538
QY 639 DSSVDGKRLLSSHDECIEHRRMLEQKIVDLEEFLENKKSSENDKQSSQEDFMESIQLC 698
Db 539 -----MDAQITAKLEKARSEISGMEKSDK---NOGCFSPKAOKE 574
QY 699 EAIMAKANALEALMRDNNDNIILENETIKRRIADLERLAKN--QETNFEELLEKEYO 757
Db 575 PCLEFA-----NTEKRLKQLOLITQELNNSKQYEFEF---KELT 609
QY 758 KEHEAQLIHEIGSLKTLIVENAMYNONLEDEFTKTLKROEIQLOLELR--RADNLO-- 814
Db 610 RKQLELESELSLOKA-----NLNLENLEA--TRACKQOEY--SOLNKIHAETLKII 659
QY 815 ---KQVNF-----DISVSKDSEKICEEITFQKOSLSDAEAVTRDAQECSEFLSENL 865
Db 660 TTPTRKAYQLHSRPVPKLSPKWSFGSL---YTQNSSITLDDIINEPVPE-----706
QY 866 ELKKEMDTSWVYQKQKASLFEKOLETERKSNKKKEADLOKE-----IQSAFNEINY 919
Db 707 -----MNOQAFPAISEELTYQEOVSALQAKIDEEHNNKLQOQHVADLKH 752
QY 920 LNLGLACKVPRDLISRYELEKRYSEFSKOLEKALEENKALENEY--TCLSEKFLPNVEEC 978
Db 753 HSTQM-----QELFSSERI-----DWTQOEELLSQLNVELEKQOEFQTKNDFLKSEVD 802
QY 979 LKNOISKASEEIMLKOE-----GEHSASISKQETIMOEOSEQILQITDEVYTHQSK 1031
Db 803 LRVYLHSADEKELSVKLEYSFKNQKEFKLSERHHVQLOLDNRLLENKELLESKAC 862
QY 1032 VQOTEEOYLEMKKHNDLFEKTYIRN--KSEADLLREMNLLKGTMESEVVK-----IA 1082
Db 863 LQDSIDNMLQETMKETIQOLSNLONFKKENETLSLSDNNLLEAEKERNNKSLQFEE 922
QY 1083 DTKHELEETTRDKQOLLEK-KYFFQAMQTFPTPLSDSLSPSKLVEGNSODPTEIENDY 1141
Db 923 DKENSSKEILKYLEAVQOEKQEKAKCEQOMAKQKLEESL-----963
QY 1142 HNLIALTERNNIMVCLTERBNSLKEBOYIDLNTQLOSLQAOISTEKSD--LOKPKODLEEGE 1200
Db 964 ----LATER--VYSLEKSSDSQKRVVADIMNOIOELRSSVCKEKTEITDILKQELDIN 1016
QY 1201 VK---LLEEMELKGLHLDISQLEKLOLENLEVTKELOTLOEBMKNTITERNLQNF 1257
Db 1017 CKYNSALVDR-----ESKVLTKQEVOLDLKEFLR-----LRLISE 1054
QY 1258 DLKAEHSLKODLSSENIEQ-----SIETODELRAQOEELREKQOLVDSFRQOLDCSV 1310
Db 1055 DI---ERMMLCEDLAHATEQOLNMLTEASKKHSGLQSNQOEBELTKKEALIOELOHNL-----1107
QY 1311 GISSPNHDAVANDKYSVLGEVNSLOSQEMLNGERDELQTSKALVSELELLRAHKVSEGE 1370
Db 1108 -----NOKK-----EEVEOKKNEYNFKMKQL-----EHYMSAAE 1137

QY 1371 N-----LEITKINCLEREKELIKSESESEVLKSMLEIKEDNNKLE--QAEYS 1417
 Db 1138 DQSPKTPPHQTHLAKLETQOEI---EDGRASKTSLEHVTAKLNEDEEVANALIR 1193
 QY 1418 SKENQFLEEVFSSQKLYDEIVLKAOLKRAERLEIKRDYFELYOTAN-----1468
 Db 1194 MKBOLREHENTRLBSQOLITENMLLOQOLDIKROKENSODNHDOOLNKEOESIKER 1253
 QY 1469 ---TNLVEGKLEPQLADHEEDSIDRSSEMEIVLGEKLEKNQYLLERLOEKELETSK 1525
 Db 1254 LAKSKIYEBMLK--MKADLEEVQSAALNKMECELCMTDEVERQYTLSSKAPQOEQJRSK 1311
 QY 1526 LEIIOKEMESVILKDLQOKLESLSEN--TILKENIDTTLKHSQTOAOLQTOEQLO 1583
 Db 1312 LEEYERERERTSQEMEMLRKQVECLAENGKLVGHQNL-----HOKIQYVRLKRENV 1365
 QY 1584 LAKMLAASDNCPIPOEKESA 1606
 Db 1366 LAETEKLRANVFLKKEKRSSES 1388

RESULT 7
 US-09-592-054-8
 ; Sequence 8, Application US/09592054
 ; Patent No. 6440684
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; APPLICANT: Finer, Jeffrey
 ; APPLICANT: Sakowicz, Roman
 ; APPLICANT: Wood, Kenneth
 ; TITLE OF INVENTION: No. 6440684el. motor proteins and methods for
 ; TITLE OF INVENTION: their use
 ; FILE REFERENCE: 1016
 ; CURRENT APPLICATION NUMBER: US/09/592, 054
 ; CURRENT FILING DATE: 2000-07-20
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 8
 ; LENGTH: 1234
 ; TYPE: PRT
 ; ORGANISM: Human
 ; US-09-592-054-8

Query Match 6.5%; Score 953; DB 4; Length 1234;
 Best Local Similarity 27.2%; Pred. No. 2.9e-39;
 Matches 333; Conservative 220; Mismatches 406; Indels 239; Gaps 42;

QY 3 EGDVAVCVVRPLQIRE--QGDQANLQWKAGNNTISQVGT--KSPNFRVFNESHSTSO 59
 Db 6 KGIPIVAVLRCPVLPKKEISEGQMCISFVPGETQV--VVGTDKSTTYDPVFDPCIEGEE 63
 QY 60 IYQEIYAVPIIRSLAQYNGTIRAYGQTSQKTYTMG-----TPNSLGIIPOAIOEYF 112
 Db 64 VFNKAVAPLIKGIFFKGYNATVLAAGQTSQKTYSMGCAVTAQENPVTGIIIPRVIOLEF 123
 QY 113 KIIQEIYNEEFILIRVSYMEIYETVVDLCCDRKKRKLIRDFRNANVVAULTIELVAY 172
 Db 124 KEIDQKSDPEFLTAKSYLTIYIEEIDLCPSREKAQINIRDPKGIKIVGLTEKTYIV 183
 QY 173 PEHVIVQWIKGKERNHYGTETKANDHSSRSHITFRMIVESRDNDPTNSENCQAVVSHL 232
 Db 184 ALDVTWSCLEQGNNSRTVASTANSSOSRSHALFTISLBEGKSSDKNS-----FRSL 236
 QY 233 NYVDLAGSRASQGTAEGRVLEGCNINSLTILGVYIKKLSQDGAQGFINTYRDSKLTFR 292
 Db 237 HLYVDLAGSRQKTRAEGRVLEGCNINSLTILGVYIKKLSQDGAQGFINTYRDSKLTFR 296
 QY 293 LQNSLGNKAKYIITCTIPV--SPDETSTTQFASAKHVRTPVNVNVLDEALAKTYR 350
 Db 297 LQDSLGNSHITLAKVSPADSNSLEETISTLYADRAKTKKPYLV-----IDPHT 348
 QY 351 KEIIDLKQLEN-----LESSETRQAQMAKE--EHTQLAEIKOLHREKREDRIWHITNI 403

Db 349 AELNHLKQYQVQLVLLLAQHGSLPQSTINAPSENLQSIEMKNQSLVEENKLSRC--L 406
 QY 404 VYASSQESQODQYKRRRYTAPKTIQNSLHSGVSDPMISRLPGNFSKAKFSDMS 463
 Db 407 SKAQGTQAMLERIILTEQVNERKLNKLEIRLOHAAKIDL-----QKIVETLEDOE 458
 QY 464 FPIIDSVV-----TEFSDFDALSMDSNGIDAEMLNLSKYTHREKTS-----507
 Db 459 LKKNVEICHLQOLITQLSDEYVACT---AAADIVVEEBAQVETSPETSRSSDAFTQH 515
 QY 508 -LHOSMI--DEGQISDSVOFHD-----SKENQL--QYLPDQSGDMAECRKAPEKEI 555
 Db 516 ALHQAOQSKREVVELNNALMLKEALVRKMTQNDQLOPQIQFOYQDNKIKNLEIVNLQK 575
 QY 556 TSLQOQLOKRE-----EKKELVQSEFLKIALLEQLSYKAKNLEMTNSRHSIN 606
 Db 576 EELVRELQTKAKNANOAKISEHRRKILLOLEQIADLKKRLKLEOSKILKL-----625
 QY 607 AEVQDVEKEVYKREMSVLDGSGYNASNDLODSVDGKRSLSSHDECIEHRKMLEQKIV 666
 Db 626 -----KESTERIVSKLNO-----EIMAKNORY- 648
 QY 667 DLEEFLENLKKSENDQKQSSQEDFMESIQLCALMAERANALELALMRNFDNITLEN 726
 Db 649 ---QLMRQKKEDEAKERQOKRKD--KEYIQKE---RDRKROYELKTLER--NQK---QS 697
 QY 727 ETLKR---EIALDERSLKEN--QETNEFELLEKEQKE---HEADLIEHIGSLKLVENA 778
 Db 698 NVLRRTEBAANAKRKLQADQKREVADKREKETSQSGNEGTAAARVMNLGIEVWVST 757
 QY 779 EYMNQNLDEDLFTKTKLKEQEIQLAEIKRADNLQKVR--NFDLSVSMG-----827
 Db 758 EBAKRLNDLDBR--KILAQDVYQLEKESRENPPIKAKCTFSLVHGVLSEDCI 816
 QY 828 -----DSKELCEIFQLQKQSLSDAENVTRDQKEC-----SLRSEN 864
 Db 817 TQOISLEFTEMLRSQIADLQOKLIDAE--SEDPKQCMENIATILKAKLAKLIGEL 874
 QY 865 LELKEKEDTSMWYNQKEKAASLFEKOLETEKSNYKKMADQOKEL-----OSAFIEIYVL 920
 Db 875 VASKIHVTKLENSLQKASQADQKMLFEQONHSELETIELQADLVYMEQDHOEKVYL 933
 QY 921 NCLLAGVPRDILLSVLELKKVSEFSKOLEKALEEKNALLENVTLSSBYKFLPNEVECK 980
 Db 934 -----LVSOLO--ESQMAE--KOLEKSASEKE-----POLVSTLQ 964
 QY 981 NQISKASEIMLLKQEGESASIIISKQELIMQOSEQILQLDDEVTHQSKVQOTEF---1037
 Db 965 CO-----DELEKAKREVCENQOOLQOENETIKOKLI--LLQVASROKHLPNTLSPDSSF 1018
 QY 1038 QYLEKKKMHDDLFEKYIRNKSEADL-----LREMENLKGTMESVE 1078
 Db 1019 EXIPPKPRPSRYKEKFLQSGMDIEDLKCSHPSVNEHEDGDDGDSDE 1066

RESULT 8
 US-09-592-054-2
 ; Sequence 2, Application US/09592054
 ; Patent No. 6440684
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; APPLICANT: Finer, Jeffrey
 ; APPLICANT: Sakowicz, Roman
 ; APPLICANT: Wood, Kenneth
 ; TITLE OF INVENTION: No. 6440684el. motor proteins and methods for
 ; TITLE OF INVENTION: their use
 ; FILE REFERENCE: 1016
 ; CURRENT APPLICATION NUMBER: US/09/592, 054
 ; CURRENT FILING DATE: 2000-07-20
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 1232

TYPE: PRT
ORGANISM: Human
US-09-592-054-2

Query Match 6.4% Score 945.5; DB 4; Length 1232;
Best Local Similarity 24.6%; Pred. No. 6.8e-39;
Matches 372; Conservative 279; Mismatches 489; Indels 369; Gaps 61;

QY 3 EGDAVAVCVRRVPIQRE--OGDOANLQWKAANNISQVDT--KSFNDRVFNHSTSQ 59
DB 6 KGIIVRALRCRPLPKPEISGCMCLSFVGEPOV--VVGTDKSFYDFEFDSTEEQE 63
QY 60 IYOBIAPIIRSAOGNGTFAVGOTSSGCTYMMG-----TPNSIGTIPOALQVVF 112
DB 64 VFNTPAVPLIKGVKGYNAVALAYGOTSGKTYSMGAYTAEOENEPVGVIPRIQLLF 123
QY 113 KIIOEINREFELVSYMEIYNETVKDLLCDRRKKPLREINEDERNRYVVDLTELMMV 172
DB 124 KEIDKSDPEFTLKVSLYEIYNEEILDLCPSRKQAQININEDPKKGIKIVGLEKTYLV 183
QY 173 PEHVYQIKKGEKRRHYGETKANDHSRSHITIFRMIVESRDRNDPNSNCGAVMSHL 232
DB 184 ALDVSCLQEGNNRRTVASTAMNSQSSRSHAILTISLEQKSKDNSS-----FRSKL 236
QY 233 NLVDLASERASQTAGSVRLKEGCNINRSLFLLGOVYTKKISDQAGGFIVYRDSKILRI 292
DB 237 HVLVDLASERQKTKAGDRLEKGININRGLLCLGNVYISALGDQKGGFADYRDSKILRL 296
QY 293 LQNSLGNAKTVIICITPV--SEDETSLTQFASAKHVNTPHVNELDDEALLKRYR 350
DB 297 LQDSLGNSTHMTACVSPADSNLEPTLNTLRADRAKIKRPVY--IDPQT----- 348
QY 351 KEIIDLKQLEENLESSESTKAQAMAKEBHQLAEITQIKHREDRIMHNTIYVASSQE 410
DB 349 AELHMLKQVOYL-----QVLL----- 365
QY 411 SQOQVRVRRKRVYMAPGKIQNSLHASGVSDFDMLSLRPGNFSKAKPSPSPPEIDS 470
DB 366 -----LQAHG-----GTLPGSITE-----PS----- 382
QY 471 VCTEFSDFDALSMQDSNGIDAEMNLASKYTHREKTS--LHOSMIDGQISDSVOFHDSS 528
DB 383 -----ENLOSMEKN-----QSLVEENKLSNGISEAAGQTAQMERIIMTQEA 426
QY 529 KENOLQTLPRDSDGMACEKRC--ASFEKETTSLOQLOQSE--BEKELVOSFEKLAELQ 586
DB 427 NE-----KNNAKLEELRCHAAKLDLOKLVETLEDELEKNEVEIICNLQOLLTQJLSE 479
QY 587 LSVAKMLMEVNTNSREHSINAEVOTDVEKEVYRKEMSVIAGDSYNASNSDLQDSSVDQKR 646
DB 480 -----IVACMAAAIDVADEAOVETSP-----TSKSSDAFTTOHALR 518
QY 647 LSSSHDECIEHRKMLEOKIYDLLEEFIEIENLANK--SENDKQ--KSSBODEMESIOLCEAIAAE 704
DB 519 QAQSKRELVELNKLALK-----FALARKMTQONDSQLPIDQYOYODNI----- 561
QY 705 KANALELALMRDNFDNIILLENETLREIADLESLEKNOETNEFEIIEKTOKEHEHQ 764
DB 562 KEPELEVINIQKEK--EELVLELQJAKDANOAKLSERRRKLQLEQIADLKKR----- 615
QY 765 IHEIGSKTLVEMNAEMNQNLEEDLETKTKLBOEIOALRKRADNLOKRVNFDLSV 824
DB 616 LNEOSKILLKESTERYVSKINDOI-----RMMKNQRYOL--MRQKKEAE--KFPQW----- 664
QY 825 SMGSEKILCEIEIFOLKOSLSDAEAVTRDAQECSEFLSENLELEKEMEDSNMTYNOKEKA 884
DB 665 -----KQKRDKEVIOLEK-----RDKRKQ-----YELKLERNPQKQSNVLRKKEE 706
QY 885 ASLEKOLETEKSYKKMEADLOKELQSAFNEIYLNGL--LAGKVPBDLLRYVLEKKV 942
DB 707 AAANKRILKDALQKQREV--ADKREKETS-----RGMESTAARVKMWMLGNEIEVWST 757
QY 943 SEFAKQLEKALEKNALENEVYTCISEYKFLPNEVECLKNQSKASEEIMLKGSEHSAS 1002

DB 758 EAKRHLNDLEDKRIKLQDVAQLKE-----KKEGGENPP 793
QY 1003 IISKOELIMOBSEQIILQDTDEVHTQSGVOOTEBOYLEMKRMHMDLEFKYIRKNSAEAD 1062
DB 794 KLRRTEFLTEVRGVSSSEDSIT--KQLESLETEMEFSAQIA--DLQOKLL--DAESD 848
QY 1063 LLRE-----MENIKGTMESVEYKIADTKHELE--TIDKQQLHEKKYF 1105
DB 849 RPKQRMENIATLEKAKLKYLIGELVASKIOVSKTLESLSLQSKTSCADQMKMLEERNH 908
QY 1106 FOAMOTIPIPTPLSDLPKSLVEGNSODPIEINDYHNLALATERNNIMVCLTERNSL 1165
DB 909 FAEIFTEL-----QAEIVRMEQHPKV-----LYLLSLOQOSOMA-----EKQ 948
QY 1166 KEQVIDLNTQLOS--LQAGSIEKSDLOKRPDLEGEVYKLLLEMLKHLTDSQLSIEKL 1224
DB 949 EESVSEKQOQLOSTLKQD---EELKMRVEQEQNO--QLRBNETIKQKLLQVYASRQK 1004
QY 1225 QLEN---LEVTEKQTLQEBKKNITIERN--ELQTNFEDLK-----AEHDSLKQDSE 1272
DB 1005 HLPKDTLLSPDSSEFYVQPKRPSRKKEKLEQSMIDIEDLKYSSEHSYNEHEDGDGDE 1064
QY 1273 NIEOSIETODELRAAOEELR-----EOKOLVDSFROQLLDCSVG--ISSPNHDAVANO 1323
DB 1065 GDDEHMKPTKLVNYSRKNIQGCSCGKMGCKNQ--GCKKQKSDCGVDCCPPTCKRNRQ 1122
QY 1324 EKVSIGEVNLSQEMNLGERDELOTSCALVSELELLRAHYKSVGENLEITTKRLNGLEK 1383
DB 1123 GKDSIGTVERQO-----DSSESK-----LE-----DPTVYTPGLSPFNP 1157
QY 1384 EILKSESEVYKSM--LENLKE-----DNKKLEQAEER-----SKENQFSL- 1425
DB 1158 --VCATPNSKILKEKMDVEQVLSKTPPAPSPFDLPDLKHVATEYQENKSGKKRRKALA 1215
QY 1426 --EEVYSGSKLVD 1438
DB 1216 SNTSFFSGCSPIEER 1230

RESULT 9
US-09-722-139-2
Sequence 2, Application US/09722139
Patent No. 635471
GENERAL INFORMATION:
APPLICANT: Beraud, Christophe
APPLICANT: Freedman, Richard
TITLE OF INVENTION: No. 6355471el motor proteins and methods for
TITLE OF INVENTION: their use
FILE REFERENCE: 1055
CURRENT APPLICATION NUMBER: US/09/722,139
CURRENT FILING DATE: 2000-11-24
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 1375
TYPE: PRT
ORGANISM: Human
US-09-722-139-2

Query Match 6.1% Score 906; DB 4; Length 1375;
Best Local Similarity 25.5%; Pred. No. 6.7e-37;
Matches 413; Conservative 247; Mismatches 531; Indels 430; Gaps 71;

QY 6 AVKVCVRRVPIQRE--LQWKAANNISQV-----DGTKSFNDRVF 51
DB 3 SVKAAVRRVRRPNRRKDEAKFIQMEKSKTITNLKIPBEGTDSGRNERTKFTYDPSF 62
QY 52 NSHETS-----QIYOEIVPIIRSALOGNGTIFAYGOTSSGKTYTMMGTPNSLGI 104
DB 63 YSADTQSPDYVQSEAVFETLCTGDYVKSFAFEGNACVFAYGOTSGSKSYTMMGNSGDSGLI 122
QY 105 PQALQVETKIIQEI--PNREFLRAVSMEIYNETVKLLODDRKKP-----LEIREDFN 157


```

Db 123 PRICEGLSRINETTRMDEASFRTEVSYLEIYNERVRL--RRKSKTENTLNRREHPK 179
QY 158 RRVYVADILEELVMPENHVIQWIKGKKNRHHGETKANDHSRSHITP--RMIVESDRN 215
Db 180 BGPVEDSKHLVQYGVVEELMDAGNINRTTAAGMNDVSRSHAITIKFTQAKFDPSE 239
QY 216 DPTNSDCGAVMWSHLNVLDSERASOTGAEGVRLKEGONINRSFLITGOVIKKLSD 275
Db 240 MP-----CE--TVSKHILVLDAGSERADANGATVRLKEGONINRSFLITGOVIKKLSD 289
QY 276 GOAGGFYIRDSKLTIRIIONSLGNAKVIITCTTP--VSFDELSTLQFASAKHVRNT 333
Db 290 -----FVEYRDSVLTWMLKDSIGNSKTIATISPADVNGETLSTLRVANKNITNK 344
QY 334 PHVNEVLDEALKRREKIEIDLK-----KOLENLESSETKQAMAKEHTOL-LAET 386
Db 345 PTINDANV-KLIRLRAETIRKTLAOGNOIALDSTP-----ALSMERKLOONEARV 398
QY 387 KOLHKEEDRIWHLTNIVASSQESQDQVYKRRRYTMAPGKIONSLHASGVSDFMLIS 446
Db 399 QELTEWNTNK-WNET-----QNTLKREOTL-----ALRKEGIG----- 429
QY 447 RLPGFSSKAKFSPMPREIDDSVCT-----ERSDDDLASMD--SNGIDA 492
Db 430 -----VIDSELPHILIGIDDLSTGITLHLKREGQTVGHDASTEDQDIYHIGIDL 481
QY 493 E-----WNLSKTYHREKTSLSHOSMIDFGQISDSVOFHDSS-----KENOLOY-LPRDS 540
Db 482 ESEHCIFENIGCVLT-LIPLSGQCSVNGVOIVATHLNGAVALLGRTNFRNHRKRA 540
QY 541 GDMACRK-----ASFEKEITSLQOOLOSKEEKEVLO-----SFEKIALEBQULSVK 590
Db 541 AKLREKRSGLISFSLSMDL-----SKRENLSAVMLXNPGLEFROOREELEKLESK 595
QY 591 AKNLEMYTNSRHSINAEOQ-TYDEKVEYRKEMSVLGDGYNASNDLODSSVGRKLSS 649
Db 596 RKLIEEM-BEOKQDKALELMQOEVEFQKRETIY-----OLO-----IR 635
QY 650 SHDCIHRKA-LEOKTVDEEFTEINLKSSSENDKSSQODEMESIOLCEAIAERANA 708
Db 636 KOESLKRSHFNKLNKDL-----LAERKEFEERLRQ----- 670
QY 709 LEELALRDNPDNITLENLTKREIADLER--SLKENQENNEPELLEK--ETOKEHBAOL 764
Db 671 -QELTELQKR-----QEEFPLRQOELOLKLLENNEKEKFOIQLDOLKEDQEOY 724
QY 765 IHEIGSLKLVENAE--MYNOLLEDELTK--TKLKEOEIOLAEIRKRDNLQKVY 817
Db 725 A-KLELEKRRLEBOEKQVNLVAHLEQRLREKQEMIQLLRGEVQWEEERK----- 775
QY 818 RNFDLSVMSDSEKLCIEITQLOKOSLS--DAVATYRQAKCSCSLRSNLELKEKEMED 873
Db 776 -----DLGIRISLLRVRKEARAGGDEDELEKALRFEFERRQOLVAKVLEKRO 825
QY 874 TSNMYNOKERASLFEKOTLEKSNYKMEADLOKELASFNINVLNGLACKVARDLL 933
Db 826 LVOOKDILKEVQPEOELFCLKCEHDKESRLLEKHDSTVD-----VTEVQODE 876
QY 934 SRVLEKAKVSESKOLEKALEENNALENYTCLSE-----YKFLPNEVDELKN--QISK 985
Db 877 KIRKPEVRLQYKEROLO-----YLLQNHLPILLEKORAFELIDGSLDITLQYOKER 930
QY 986 ASEEMILKQEGHSASISIKOELIMQOSE-----QILOLTDVY--HTOSKVQOTEE 1037
Db 931 EME-----KEROLAOYQANANOLOKLOLATEEFTNINARQOEKAKKKEK 974
QY 1038 QYLEMKRMHDLFEKYIRNKSSEADLLREMEKNGTMESEVVIADTKAE--LEETIROE 1096
Db 975 EILES-----REKQOREALERALARERHSLQ-----RSTIGTELEEBOR 1016
QY 1097 QLLHEKKYFFQAMQITFFITPLSDSLPSKLYEGNSODPIEINDVHNILATERNINIV 1156

```

```

Db 1017 Q-----KIASLNSG-----SREQSGLOA 1034
QY 1157 CLETERNSLKQVYIDLNTLOSLQASIERKSDLOKPRODEEGEV-KLLEMEELKGLHT 1215
Db 1035 SLEAFQALEKQDQRLREYELQLOKQIYEVQVQKDHGHTLEGVASSSLPVAEKSHLV 1094
QY 1216 -----DSQS--TEKQLENLEYTELQTL-----QEMKNITIERNELO 1253
Db 1095 PLMDARINAYIEE-----EVQRLODLHARVISEGCSADTKMKDKLNGITQRLKY 1148
QY 1254 TNEFDLK-----AEHSLKODLESENIEOSIETODEL-----RAQOEELREOK 1295
Db 1149 ELCDLILCYLMPERDAACANHPLQODL--VOLSDMKTEIPDLVLPNGVYSSKFOQ 1205
QY 1296 QLYDSFRQOILLDSQVGISSPNHDAVANOKEKVSIGEVNSIOSEMURGERDELQTSKALVS 1355
Db 1206 TLVDMI--YFLHONMEVNP-----SLAEVOLLTYTVKMGDSGHQCCQSLV- 1251
QY 1356 ELETTRAHVSVE-----GENLEITKRLNGLEK-----EILKSSSESY 1394
Db 1252 ---LNLNHALVYEDCYFYRIRSRNIPPGAOFDVY-KCHALSERCVYVPEKKNVSTV 1307
QY 1395 LKSMLEMLKEDNNKLKQAEYSSKEN--QFSLEEVFSQOKLYDBIEYLKQILKAER 1452
Db 1308 ELVFLQKLPKPSVGSRSNSPPEHLQEPNVOLFTPLYLQSGQVAP--EYWKILTFNSQDA 1365
QY 1453 L 1453
Db 1366 L 1366

```

```

RESULT 10
US-09-721-832-2
; Sequence 2. Application us/09721832
; Patent No. 639346
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. 639346el motor proteins and methods for
; FILE OF INVENTION: their use
; FILE REFERENCE: 1055
; CURRENT APPLICATION NUMBER: US/09/721,832
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: Human
US-09-721-832-2

```

Query Match 6.1%; Score 906; DB 4; Length 1375;
 Best Local Similarity 25.5%; Pred. No. 6,7e-37;
 Matches 413; Conservative 247; Mismatches 531; Indels 430; Gaps 71;

```

QY 6 AYKVCYRVPPLQREGDOAN--LQWKAGNNTISQV-----DGTKSFNFRVFP 51
Db 3 SVKAVARVAPNMRKREKDEAKFLIQEMKSKTITIMLKIPGSGTGDGSGREHRTFTYDESF 62
QY 52 NSHESTS-----OIYOEIAVPIIRLSALOGYNGTIFAPGOTSSKTYTMMGTMSLIGIT 104
Db 63 YSADTKSPDYVQSEAVNFTLIGDVYKSAFEGYNACVFAYGOTGSGSKTYTMMGNSGDSGLT 122
QY 105 POAIOEVRKIIOEI--PNREFLRVSYMEIYNETVKDILCDDRKKRP-----LEIREDFN 157
Db 123 PRICEGLSRINETTRMDEASFRTEVSYLEIYNERVRL--RRKSKTENTLNRREHPK 179
QY 158 RRVYVADILEELVMPENHVIQWIKGKKNRHHGETKANDHSRSHITP--RMIVESDRN 215
Db 180 BGPVEDSKHLVQYGVVEELMDAGNINRTTAAGMNDVSRSHAITIKFTQAKFDPSE 239
QY 216 DPTNSDCGAVMWSHLNVLDSERASOTGAEGVRLKEGONINRSFLITGOVIKKLSD 275

```

```

Db 240 MP-----CE---TVSKIHLYDLAGSERADATGATGVRLKEGNINKSLVTLGAKKKQV-- 289
QY 276 GAGGFINRDSKLRILNLSIGNAKTYIICTIP--VSPDETSTLOFASTAKHVRT 333
Db 290 -----FVPRDSVLTWMLKDSLGNSKTMITATISPADNYGETLSTLYANRAKNIINIK 344
QY 334 PHVNEVLDEALLKRYRKEIIDLK-----KOLENLESSSEETKAQMAKEEHTOL-LAEI 386
Db 345 PINEDANVK-LIRELRAIARLKTLLAOGNOIALDSPT-----ALSMEERLOONREAV 398
QY 387 KOLHREDRIRWHLNITIVASSQESQODQVRKRRTVAPAKIONSILHAGSVSDMLS 446
Db 399 OELTKEWTKR-WNET-----ONILKEQTL-----ALRKEIG----- 429
QY 447 RLPGNFKRAKFSMDPSFEIIDSVCY-----EFSDFDALSMMD--SNGIDA 492
Db 430 -----VLDSELPRLHIGIDDLSTGIIYHLKEGQTYVGRDASTGEOIIVLHGLDL 481
QY 493 E-----NMLASKVTHREKTSLHOSMIDGQISDSVQFHDSS-----KENOLQY-LPKDS 540
Db 482 ESEHCIFENIGTVI-LIFLSGSGQSVNGVQIYEATHLNOGAVILLGRIMRFRNPKRA 540
QY 541 GMAECRC--ASFEKETITSLOOQLOSKKEEKKELVQ-----SEELKIALEBQOLSVK 590
Db 541 AKLREKRKSGLLSSFSLSMTDL-----SKSRNLSAVMLYNPGLPEPQORELEKLESK 595
QY 591 AKNLEAVTNSRHSINAEVQ-TDVEKEVYRKMSVYLGDSGYANSDLODSSVDGRKLS 649
Db 596 RLILEEM-EKOKSDKAELERMOQOEYEQRKETEYV-----QLO-----IR 635
QY 650 SHDECIEHRKM-LEOKIVYLEFIEIENLKNKSENDKOKSSEODFMESIQCEAIVMAEKANA 708
Db 636 KQEBESLKRSHIENKIDL-----LAKEKFEERLAEQ----- 670
QY 709 LEEALMRDNFDNIILENTLKEITLADLER--SLKENQETNEFEILEK--ETQKEHAOL 764
Db 671 -OEIIELOKKR-----QEEETFLVQOEILORLKELNNEKAEKFOJEOILOQKEDQY 724
QY 765 IHEIGSLKLVENAE-----MYNONLEEDLETK--TKILKBOEIOALERKADNLOKKV 817
Db 725 A-KLELEKRLKEOEBOVMYLAHLEBQUREKQEMIQLRKEGVQVVEEKR----- 775
QY 818 RNFDSVSGSEKCEETIOLQKSLSS-----DAEAVTRDAQECSEFLSENLEKEMED 873
Db 776 -----DLEGIRRESLRYKVARAGDEGELEKAKOLFFERKROQYKVLVLEMD 825
QY 874 TSNMTNQEKAASLEFKEOLETEKSNYKMEADLOKELQSAFNEINYLNGLAGVPRDIL 933
Db 826 LVQOKDILKKEVQOEIIECLCKCEHDKESRLLEKHDESVTD-----VTEVQDRE 876
QY 934 SVELEKQVSESKOLEKALEKKNLENVYGLSE-----YKFLPNEVECLKN--QISK 985
Db 877 KIKPEYRQYKERQLO-----YLLQNHLPYLBEEKORAFELIDRGLSLDNTLYOVEX 930
QY 986 ASEETMLKQEGEHSASISKOIIMQOSE-----QIOLQTEEVT-----HTOSKYQOETE 1037
Db 931 EHEE-----KEEQLOAQYANANOLOKLOLATEEFLANINAKROEKYKAKKEX 974
QY 1038 QYLEKKNHMDLFEKYIRNKSAEDLLREMNULKGTMESVEYK IADTKHE-LEETIRDE 1096
Db 975 EILLES-----REKQOREALERLARLERRHSALQ-----RHSGLTEIEEGR 1016
QY 1097 OLLHKKYFFQAMQITPITPILSDSLPRSKLYEGNSQDIEIENDYHNLALATERNNIIV 1156
Db 1017 Q-----KLASLNSG-----SREOSGLQA 1034
QY 1157 CLETERNSLKEQVINDLNTQLOSLQASILEKSDLOKPRKDLGEGV-KLLLEMLLKGHLT 1215
Db 1035 SLEADQEALEKQOERLEIIOOLKQIYEVGDVQKDHGTGLGKYAASSLSPVSAKSHLY 1094
QY 1216 ---DSQLS--TEKLOLENLEVTEKLOTL-----QEEKNKTTIERNEIQ 1253
Db 1095 PLMDARINAYIEE-----EVOORLODDHRYISEGOSTSADTMKDNKELHNGTIOIRKLY 1148

```

```

QY 1254 TNFEDLK-----AEHDSLKODISENIEOSIETODEL-----RAAOELERQK 1295
Db 1149 ELCRDLCLVMPPEPPAACAANHLQOOL--VQSLDMKTEIPDLVLENGVASSKFT 1205
QY 1296 OLVDSEFROQLLDCSVGISSPNHDVANQEKESIGEYNSLOSLEMLNGERDELOTSKALS 1355
Db 1206 TLVDMT--YFLNGMMEVNP-----SLAEVOLLITTVYMGDSGDHDOGSOLV- 1251
QY 1356 EELERAHVSYE-----GENLEITTKLNGLEK--ELIKSESESV 1394
Db 1252 ---LMTHTHALYKEDCFYPRIRSRNI PPPGAOFDVI-KCHALSEFRCVVPEKKNVSTV 1307
QY 1395 LKSMLENTKEDNKKLEQAEYESKREN--QFSLEEVFSGSKLYDIEIYLKQOLKAAER 1452
Db 1308 ELVLOKLPVSGSRNSPPEHLOEADPVOLFPTPLYLQGSQVAP--EVMKLTENSQDA 1365
QY 1453 L 1453
Db 1366 L 1366

RESULT 11
US-09-721-689-2
; Sequence 2, Application US/09721689
; Patent No. 6440685
; GENERAL INFORMATION:
; APPLICANT: Berard, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. 6440685el motor proteins and methods for
; FILE REFERENCE: 1055
; CURRENT APPLICATION NUMBER: US/09/721,689
; CURRENT FILING DATE: 2000-11-24
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FASTSEQ for Windows version 4.0
; SEQ ID NO 2
; LENGTH: 1375
; TYPE: prt
; ORGANISM: Human
US-09-721-689-2

Query Match 6.1%; Score 906; DB 4; Length 1375;
Best Local Similarity 25.5%; Pred. No. 6,7e-37;
Matches 413; Conservative 247; Mismatches 531; Indels 430; Gaps 71;

QY 6 AVKCVVRVRLQIREGQGAN--LQWKAGNNTISQV-----DGTKSFNFRVFE 51
Db 3 SVKVAVRVPRMNRREKDLAKFLIQMEKSKTITNLKIPGGTGDSGRERTKFTYDFSG 62
QY 52 NSHESTS-----QIYQEIAPPIIRSAQNGITFAFGQSSGKTYTMAGTSPNSLGI 104
Db 63 YSADTKSPDYVSQEMVFKLGTDVYKSAFEGYNACVFAFGQSGSGSYTMGNSGDSGLI 122
QY 105 POAIOVEFKIIOEI--PRREFLRYSYMEIYNETYKDLCDORRRKP--LEIREDEN 157
Db 123 PRICGELFRINETRMDASFRTEVSYLEITNERRDL--RRKSSFTMLRPREHKK 179
QY 158 RNVYVAADLTELVMPEVHIQWIKKGEKNRHYGETKMDHSSRSHTF--RMIVSRDN 215
Db 180 EGPYVEDLSKHLVYQVGVDEELMDAGNINRTAATGMNDVSSRSHAIFIKFTQAKFDE 239
QY 216 DPTNSDCGAAVWSHLNLYDLAGSRAQTAEGVRLKEGNINRSFLITLOVIRKLSLSD 275
Db 240 MP-----CE---TVSKIHLYDLAGSERADATGATGVRLKEGNINKSLVTLGAKKKQV-- 289
QY 276 GAGGFINRDSKLRILNLSIGNAKTYIICTIP--VSPDETSTLOFASTAKHVRT 333
Db 290 -----FVPRDSVLTWMLKDSLGNSKTMITATISPADNYGETLSTLYANRAKNIINIK 344
QY 334 PHVNEVLDEALLKRYRKEIIDLK-----KOLENLESSSEETKAQMAKEEHTOL-LAEI 386
Db 345 PINEDANVK-LIRELRAIARLKTLLAOGNOIALDSPT-----ALSMEERLOONREAV 398

```

```

QY 387 KOLHREDRILWHLTIYVASSQESQODQRYKRRVTWAPKIQNSJHASSVDFDMLS 446
DB 399 QELTKEWINK-WNET-----QNLKEQTL-----ALKEGIC----- 429
QY 447 RLPGNFSKAKFSDMPSPEDDVSCT-----EESFDALSMND--SNGIDA 492
DB 430 -----VVLDELPHIGIDDDLLSTGIIYLHLKEGQYVGRDASTEDQIYVHGLDL 481
QY 493 E-----WNLASKYTHREKTSJHOSMIDFGQISDVQYHDS--KENOLQY-LPBD 540
DB 482 ESEHCFIFENIGTYT-LPLSGSQSVNGVOIVTEATHLNGQAVIILGRTNMFRENHPEKA 540
QY 541 GDMACRCR--ASFKEITSLQOQLOSKREKKEKELVO-----SFEKIAELEROLSVK 590
DB 541 AKIRREKRSGLTSPSLMTDL-----SKSRENLSAVMLYNGLEFEGQREBELEKESK 595
QY 591 AKNLEMYNRSRHSINAAYQ-TDVEKEYVRKEMSVLGDGYNASNSDLQDSSVDGKRLLS 649
DB 596 RKLIEEM-EEROKSDKAELERMOQEVETORKELEIV-----OLQ-----IR 635
QY 650 SHDECIEHRKM-LEQIYDLEEFLENLKNKSENDKOKSSODPMESTQLCAIMAKKANA 708
DB 636 KOESLAKRSHFHEIKLKL-----LAKKEFPERERLEQ----- 670
QY 709 LEBLALMDFNDNITLENETLKREIADIER--SIKENQETNEFELIEK--ETOREHEAOL 764
DB 671 -OEIETLOKR-----OEFEETFLVQOEELORKEILNNKAKKFOIFOLDLOKEKEDQY 724
QY 765 IHEISLKKLVENAE---MYNONLEEDLETK---TKLKQEOLOAMELRKANLOKRY 817
DB 725 A-KLELEKRLLEBQEKQEVYVAHLEBOLREKQEMQIOLRGEVQWVEEER----- 775
QY 818 RNFPLSVSMGDSKCEIEFQLOKSLS---DABAVTRDAQKESQFSENLELEKEMED 873
DB 776 -----DLEGIRSELILAVKEARAGGEDGELEKAQURFEFRQYVRLKVLVNEKD 825
QY 874 TSNMYNOKERAKASLEKQLEFEKSNYKMKEMADLOKELOSAPNEIYINGLAGVYRDL 933
DB 826 LVQOKDILKKEVQOEQELTECLCEHKEHESRLKEHDSVYD-----VFEVPODFE 876
QY 934 SRVELEKKEVSEFSKQLEKALEKNALENEVCLSE---YKFLPNEVECKLN---OLSK 985
DB 877 KIRVVEYRLQYKERQLO-----YLLONHLPTELEKQRAFEIILDRPLSLDNTLYVEK 930
QY 986 ASEIMILKQEGHSASIIKQETIMOQOSE---QILOLTDDEVY---HTOSKVOQTEE 1037
DB 931 EMEB-----KEEQLOAQYQANANOLKLOKQATFEFTANIAHGOEKVRRKEK 974
QY 1038 QYLEMKKMHDDLFEKYIRKNSAEMEDLLREMBENLKGTMESEYVKIADTKHE-LEETINDKE 1096
DB 975 EILES-----REKQORALERALARLERHRSALO-----RHSITLGTIEIEOR 1016
QY 1097 QLHKKRYFQOMQITPITPLSPSLKLVESQDPIEINDYHNLILALATERNNIMV 1156
DB 1017 Q-----KLASLNSG-----SREQSLOA 1034
QY 1157 CLETERNSLKEOYIDLTFQLOSLQASIEKSDLOKPKODELEGEV-KLLEMEILLKGHLT 1215
DB 1035 SLBAEOBALKEDEKLELEYEQLOKQIYEVGYQKDHGHTLEGKVASSSLPVAESKHLV 1094
QY 1216 ---DSOLS---IEKLOLENLEVTETLOTL-----QEMENITIERELO 1253
DB 1095 PLMDAKINAYIEE---EVQORLODLHRYVISEGGSADTKMKDEKLNKTIOKRLKY 1148
QY 1254 TNEFDLK-----AEHDSLKODLSENIEOSIETODEL-----RAAEELREOK 1295
DB 1149 ELCRDLICVLMPEPDAACANHPLOODL---VOLSLDMKTEITPDLVLPNGVYSSKFOI 1205
QY 1296 QLVDSFRQOOLLDCSVGISSNHDAVANOKEKVSJGEVNSIOSEMLRGERELQTSKALVS 1355
DB 1206 TLVDMT--YFLHGNMEVNV-----SLAEVOLLITTYTVKMGDSGHQDCQSIV- 1251

```

```

QY 1356 ELELBAHVKSVE-----GENLEITKINGLER-----ELIKSESSEV 1394
DB 1252 ---LUNTHIALVKEDCYFPYPIRISRNIPPEGAOPDVI-KCHALSEFCVYVPEKKNVSTV 1307
QY 1395 LKSMLENLKEDNNKLEQAEYSKSEN--OFSLSEEVSGSQIYDEVEVLKQALKAER 1452
DB 1308 ELVFLQKLRKPSVGRSNRSPPEHLQAPVQLETTPLVYLOSGQNVAP--EYVKLTFRNSODEA 1365
QY 1453 L 1453
DB 1366 L 1366

```

```

RESULT 12
US-09-134-001C-3159
; Sequence 3159, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: CTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3159
; LENGTH: 10182
; TYPE: PR
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3159

```

```

Query Match          5.98; Score 866; DB 4; Length 10182;
Best Local Similarity 18.6%; Pred. No. 6,3e-34;
Matches 661; Conservative 697; Mismatches 1295; Indels 898; Gaps 157;

```

```

QY 15 FLOREGGDQANLOWK-----AGNN-----TISOYDGTGSPFNDRY-----FNH 54
DB 5623 PINKNEIEBQAINQINTQYALSGENKLRHDOESTNRQLEGLSLTAQIINAEDLVNQA 5682
QY 55 ESTSQIYQEI-AVPIIRSALOGNGTTFAYGQTSKRYTMMGTGPSLGIIPAIOIEVR 113
DB 5683 KTRTDVAQKLAAKEINSAMSNLRDGIQNKEDIKRSSAV-IANDPKRYAYAOALQANAN 5741
QY 114 IIOEIPRER-----LIRV-----SYMEIYN-----EYKDL-LCDDRKKPL 150
DB 5742 IINATPVVELNKATTIEQALSRYQAOQDIDGVOQLANAKQATQTVNGINSILNDGOKREL 5801
QY 151 -----EIREDFNRNVYADLTLELVMPBEHVIQ-----WIKKGENRHYGTER 193
DB 5802 NLLINSANTRTKQOEELNKATELNHAMEALRNSVQVADVQKSSNYVEDQPEQHNYDNA 5861
QY 194 MND-----HSSRSHITFRMIVESRDRNDPTNSECDGAVVSHLNLVDLAGESASQTA 248
DB 5862 VNEAQATINNNAOPVLDKLAIEHRLTQTVNTYTKDALHGA-----QKLTQOQAAETGI 5913
QY 249 EGYR-----LKEGCIINSPLTIGOVYIKKLSDGQ-AGGF 281
DB 5914 RGLTSLNEBQKAEVAKVTAATTRDEVRNIRQEAATTLTAMGLRKSISDKNDTKNSKRY 5973
QY 282 INVRDSKLTRILQNSLGNAGKATYIITTPVSPDETLSTL-----QFASAKAHVRNPH 335
DB 5974 IN-EDHDQOQADNAN-NNAAQVYI-----DETQATLSDDTINQLANAVTQAKSNLH 6022
QY 336 VNEVLDEALLRKYREKILDKOLENLESSESTEQAKAAKEEHRO----- 381
DB 6023 GPTKLOHD---KDSAKQIT--AQLONLNSAQKHEDSLIDNESTRTOYQHDLTEAQALD 6076
QY 382 -LLAEIKOLHKEHREDRILWHLTIYV-----ASSQESQO-----DQRYKR 419

```

Db 6077 GLMGLAKESIKD-----YTINIVSNGNYINAEPSKQOAYDAVQAONNIINGTNOPTIN 6129
 Qy 420 KRRTYMAPGKTONSLH-----SGVSPDMLSRLPGNFSSKAK---FSDMSEF 464
 Db 6130 KGNVTTAQTQYVKNKTDALDGDHRLLEAKNMANQTRNLISNL--NNAOKAEKNLVASATL 6188
 Qy 465 PEIDDSVCTEERDDPDLAS-----MMDNSCIDAEWNLASVYT 501
 Db 6189 EOYQONLOT-AOLDNMGELROSIACKDYKASKYINEDPOIKONTDYDAVOREYETIIN 6247
 Qy 502 HREKTSIHSOMIDFGOISDSVO-----FHDSKKNQ--LOYLPKDSGDMACRKAFFEKEI 555
 Db 6248 ETQPELLEKANID--QATQSVQNAEQALHGAEKLNQDKQTSSTELDGLTDLDAQREK-- 6303
 Qy 556 TSLQOOL---QSKKEEKELVOSFELKTA--ELEBQLSVK-----590
 Db 6304 --LREQINTNSRDIKOKIEQAKALNDAMKLEQOVAKQGVHANSPTYNEDSAOKDAY 6361
 Qy 591 ---AKNIE-AWTSNREHSINAEVOTDYEKEVYVRKEMSVLGD-----SGYNA 632
 Db 6362 NNALKQEDILNNSNPNLNQDITLNLNINIKQADNLHGAKQLOQDKNTTNOALGINLH 6421
 Qy 633 SNSDLQDSYVDGKRLLSSHDECIEHRKALEQIYDLEEFITLNLKKSSENDKQSSQEDPFM 692
 Db 6422 LMQPQKALIGAINGATSROQVAKIEKAEAA---LDEAMQLEDOQVQDDQISNPSFI 6477
 Qy 693 ESIOLECIAMAKANALEELALMRDNDNITILENTL--KREIADLENSLKE-----742
 Db 6478 NEDSDKQTYNDKIOAKKEIINO-----TSNPTLDKOKIADTLOINIKAVANNLHGQ 6529
 Qy 743 -----NOETNEFELIEKETQREHQAOLIEHIGSLKIVENAEWYNOMLLEEDLETKT 793
 Db 6530 KLAOSKODANQNLHNDLDTLE--QKNHFKPLINNAOTDEYVQKOLEIKQ--LNGDMSTLH 6587
 Qy 794 KLEKEOETOLAEELRK--RADNLOKRVNPNFLSVSNG-----DSEKICEITFOLK 840
 Db 6588 KVINDKQ-QIOHLSNYINADN--DKQWYDNAIKEAEDLIHHPDITLHKLQADLID 6644
 Qy 841 QSLSDAEAVNR-----DAQKECSFLSENLEKEKEDTISNWTNOKKASLEPEKOLE- 893
 Db 6645 QAHNELNGESRFPKALDNLALNDIDSLNSLNPQOTVANDINNHVTTLESIAQELQAKEL 6704
 Qy 894 -----TEKSNYKMEADLOKELOSASFENEY-----919
 Db 6705 NDAMKARBDISTMNOEOTIRKSNYTNEDLAQONAYNHANDKINHIIIGEDNATMDPOIILKA 6764
 Qy 920 -----LNGILAGVPRDLISRYELEKYSSEFS-----KOLEKALEKNALENEVTC 965
 Db 6765 TODINTAINGINGDOKLQD--AKTDAKQOITNFTGLTEPQKQALENITNOOTSRANVAKQ 6822
 Qy 966 LSEKKELPNEVECKKNQISKASEIIMLLKQGEHSASISQOELIMQ--ESEQILOLQD 1023
 Db 6823 LSHAKFLNGKKEELKAVAKAS---LVRONSNYINEDVSEKAEVQOIAAGQOETINSEN 6878
 Qy 1024 EYVATQSKVOOTEQYUYLEMK--MHDDLEFKYIRNK--SEAEIDLK--EMENLKG-----1072
 Db 6879 NPTISSDINTIOEINDAEQNLHG--NKLQAOQIAINEIQLONLQDLSAQOTK 6931
 Qy 1073 -----TWESVEVKIADTK-----HELEETIRDEQOLLHEKKYF-----1105
 Db 6932 LIODIGRTTTPKPAOTKLEEKAQINQAMQLOKSIOADKALINSNYLNEDSEKKLAVDN 6991
 Qy 1106 --PQAMQTIPTPLSDSLPSKLVESQDPIELND--YHMLILATKR--NNIMVCLFET 1160
 Db 6992 AVSAQEDL--INQNDPTMDISNIQATQKVIOAKDSLHANKLAQNASNLINOST 7048
 Qy 1161 ERNSLKEQVID-----LMTQLOSLQASIEKSDL--QKPKQDLE 1198
 Db 7049 NUNKQKQALNDLNLHNAQTKQOVAEITIAQANKLINNEKQTLITYEBSQNVNQOSKYINED 7108
 Qy 1199 GEVYKLLLEMLLKGH-----LTSQLSQTEKL--OLENLEVT-----EKLQLOLEKKN 1244
 Db 7109 POYONIVNDSIQGREILNGTTDDVLNNNKIADAIQNIHILTKNDLHGQKLOKQAOQDATN 7168

Qy 1245 ITIERNEL-QTNFEDLKAENHSL-----KODISENIEOSIETODELRAAQEEL---RROK 1295
 Db 7169 ---ELNVLTLNNSQROSEHEDINSAPRTEYSNDLHAKALNEAMKOLEVALENSVK 7225
 Qy 1296 QLVDSFR-----QOILDSCVGISSPNDHAAVANOEKVSLGEVNSIQSEMLRGE 1342
 Db 7226 KLSDFINDEDAQONENYNAQKAKDIINGVPSSTLDKATIED--ALLEQONAR--ESLNGE 7282
 Qy 1343 RDELQTSKALVSEIELLRAHVKSVEGENLETYKTLNG---LEKELLGKSESESEVLSM 1398
 Db 7283 Q-KLOEAKQNAVAEIDNLOA-----LNGQVLAEXTLVNOASTREVOEA 7326
 Qy 1399 LENLKEDNNKLEQAEESYSENQPSLEEVF---SGSQKLVDELEVYKAQOLKAAEERLE 1454
 Db 7327 LQKAELEAKKALTEITENKE-QIKASRFRVNASDGLQAVNYSALNGSOIATQOPE 7385
 Qy 1455 IKDRYFELVQTAINTLVEGKLEPFLQADHEDSIDRSESEMEIKVLG--EKLERNQYLE 1513
 Db 7386 LNKQVYNATQTIKT-----AENNLNGSKLALEAKSDQNSQSTIEHLOGLTQ 7430
 Qy 1514 RLQEBKELSNKLETLQ-----KEMETSVLLKDDLOQLKLSLSENIILAKENIDTT 1564
 Db 7431 SOKDKOHDILNOAQTKQOQVDDIVNNSKQDLSM---NOLQOIV-----NNDNT 7475
 Qy 1565 LKHHSQ-----TOAQOLKQOELQAKNLALIASDNCPITTOKEKISADCVHPL---EKKIL 1617
 Db 7476 VKQNSDFINEDSQDQAVNNAHIOAKKDLITA-----HPTINDKQID 7517
 Qy 1618 LITTELHQNTNEQELKEHEKLEQOVELCEVEHLLKMSIESKSSLESIOHEKHQD--E 1676
 Db 7518 QAIENIKALND---LHGSNKLSEDK-----KASEQOQNLNLSLNGOKDITIL 7562
 Qy 1677 QOLLALKQOMOVYTO--EKKELQOTHEHLTAEVDLKKNIELGLFKNEAQOKTKKEQCL 1734
 Db 7563 NHIESAPFRSQVGEKIASAKOLNMTMKALRDSIDANNIILQSSKFYNEDSEQONAYNOA- 7621
 Qy 1735 LENKEKELEQSOHRQOCELEPELMSLKDKESALETLKESQOYVYNNOEMEWMLMELEMKL 1794
 Db 7622 VNKAKNITNDQTPYMANDELOQSVLNEVKQTKDNL--HGDQKLANDKTDQATLNALNLYN 7680
 Qy 1795 NSORTVI---AERDQLODLRESEVMSLETODDLKQAEAL-----QOQKDYQOELTSQ 1845
 Db 7681 QAQNGNLETKYQONSREQVQYVLANQNLDMAKKLDDALTGDAIQTSNYINEDTSQ 7740
 Qy 1846 -----ISVLOEKISLEENOMLYNAVYKETSLE--RDDLN-----OSKOHLESE 1887
 Db 7741 QVNEDEYTDRCNKITYAEQTN--PNNASPTNINTIADKITEARKNDLHGVOKLQAOQOSINT 7798
 Qy 1888 IETLSLSEKKEPFLQOAEKXKADARKTIDITEKISNIEBOLLOQANLKLKETTERESL 1947
 Db 7799 INOMT-----GLNQAKQKQDLN--QEIQOTQTRSEV--HOVINKAKALDSMTMLTQOS 7846
 Qy 1948 IQCKEOLALNTEHLRETLKSK---DLALGKMEQERDEANFVIALTEKMSLE--EQINEN 2003
 Db 7847 ITDEHEVQOTSNYINETYNGQTAVNNAVDRKQOIIINGSNPT-----ANPLEVERAASN 7900
 Qy 2004 VTTLKEG--EGEKETFYLDQPSKOSSSQMEELRESLTKYDLOLEBAKEEISEAT--NEIK 2060
 Db 7901 VKISDIALHGERE--LNDKNSKTRFAVNLNDNLNA-----QKEALTHEICQATVISOVN 7953
 Qy 2061 NLTKAISLEEEIILQNASILNEAVSERENLRHSKQOLVSELOSLTKLSKSDH-----2114
 Db 7954 NIYNKAKALNDMKK-----LNDIYAQOQVNOQSNNTYINEDSTPQNMVYDITINHQOSIIDQ 8009
 Qy 2115 PQOSRREKDEAVNKIASIAEIKILTKEMDEFRDSK-----SLOQSSHLSSEELCTY 2167
 Db 8010 VANPTMSHDELEMANINNIKAIHMLNDGE--HKLQAKKEVANLILNSLNDLNPQDAIRL 8068
 Qy 2168 KTELQMLKQKQEDINN--KLAEKYKENDLLOHLSLKEQLODQ---QNELNREKIRN 2220
 Db 8069 VNEAQOTREKVAEJOLOSQAOLNDAMKHLNRNOSTONOSVROESKRYINASDAKREOYNHAAVE 8128

```

QY 2221 YE--LCERMDIMEKESVLRMQNEPQOEDVDARMDILSRNOEIOEMEKISAVYSE 2278
DB 8129 VENINQHPFLDKEL--IKOLTDGVNQNANDL--NGVELLDADKNANOSIPTLMLHQA 8185
QY 2279 QHTLLSSLSSELOKTEAH-----KHCMLNKESL-----SSTLSRSGFQOT 2321
DB 8186 QONALNEKINNAVTEFEVAALIGQAKLLDHAMENLEESIKKEQYKQSSNYINEDSDVOE 8245
QY 2322 EHVKLNTLOLTLNFKKVVYRTAAVKE--DHSLL--IKDYEKDLAAROKHDELRL--OLQ 2375
DB 8246 TYDNAVDHVTLELN--QVYVNTPLSTEDLEHAI NEVNAOKKOLRGOKIYOTIIDLADKELS 8303
QY 2376 CLEOHGRKWSASAEELKFCF-----IEFLNELLFRKANITIOS--V 2414
DB 8304 KLDDLTSSOSSSSISNOITACTRTREVAQAIEKASLNAHMAKLNKVVYKNADKVLSSRFI 8363
QY 2415 ODDSEYOVF--LNQVVS-----TLOEELHKGFMQMLEEFDLHVDAKL 2459
DB 8364 NEDPEKKAAYOALNHDSIITHROTNPMDPTVINSTHETAO--NNLHGD--OKL 8417
QY 2460 SEGMOQENRIASTIOLLTKRLKAVOSKIOREITVYLNQPEAKLOEKKKEKELMRME 2519
DB 8418 AHAQODAAVYNGLIHL-----NVAQREVMINTNT--NATTRKAVAKNLNDNAQALD 8466
QY 2520 HHGFSASVMEENARILGILTYODESKKLOSRIKML--FENELNVKDDAMHKEKVAI 2576
DB 8467 KAMELTOOVVAHKNNLINDSKYLNEDSKYQOQYDRIADABOLLNQTNPYL--BPKYVDI 8525
QY 2577 LODRLLSNAAEELNAMOVLTKKODNLQAA--MKEIENLOKMAKGAOPY--KEEIDNL 2632
DB 8526 VKDNVLA--NEKILFGAELSTYDKSNANDEIKHMVYLNNAQOSIKDMIASHALTEYKOL 8584
QY 2633 -----KTKVVKIEMEKIKYSKAT-----DOEIVALKSCLEDEKESG-- 2667
DB 8585 LQQAARIIDEAKMSLEDKTYOYITDTLPTNTYASEDKKERYDVQVSHQAATI--DKINGSN 8643
QY 2668 -----LARIKEELRRA--QADNONTYCVPKYOXASTPPTVYCGGSGVSTALVLOSE 2720
DB 8644 VSLDOVQALOEQLTOASBNLDGDRVEBAKYHANOT-----IDOTLHLSLQOO 8692
QY 2721 KAALERELSHYKKRYHMLSRMTSSSEDRKTKAKSDAHSHTGSSHGSPK----- 2772
DB 8693 TA--KESVKNATKLEELIATVYNNNAQALNKVYMGLEDFINHADVEHSDNTRADDKTI 8749
QY 2773 --TETTYHGPPTPERSEMPSLHLGSPKSSSTKRVVSPNSETIYLSOLVSPGKTGM-- 2828
DB 8750 AYDALEHGO-----DIQKTNAQONETKOAL--QOLYAEFTSLNGFEELNHAR 8795
QY 2829 ----KHILSPKVLGHKKRALSPRSE-----MPTOHVISPCKTGLAHNLTESTLFDLSS 2880
DB 8796 PRALEYIKSLIEKINNQAOKSALEDKVTYOSHDLLEHETVNEG----- 8836
QY 2881 PCQOQKVOEMLNSPKGLFVYKSKSMPCPSOF-----FDNSKLGDFSEL--NTASND 2932
DB 8837 -----TJNLDMGELANAVNN--YAPTASINYINDNLKDNFTQAINNARALN 8886
QY 2933 KSOAENMWYEA 2943
DB 8887 KTGQGNLDFNA 8897

```

RESULT 13

```

US-09-724-517-2
Sequence 2, Application US/09724517
Patent No. 6379941
GENERAL INFORMATION:
APPLICANT: Beraud, Christophe
Applicant: Freedman, Richard
TITLE OF INVENTION: No. 6379941el motor proteins and methods for
FILE REFERENCE: 1031
CURRENT APPLICATION NUMBER: US/09/724, 517
CURRENT FILING DATE: 2000-11-27

```

```

PRIOR APPLICATION NUMBER: US/09/641,807
PRIOR FILING DATE: 2000-08-17
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 1279
TYPE: PRT
ORGANISM: Human
FEATURE:
NAME/KEY: VARIANT
LOCATION: (409)...(436)
OTHER INFORMATION: Xaa = any amino acid
US-09-724-517-2

Query Match
Best Local Similarity 24.7%; Score 841.5; DB 4; Length 1279;
Matches 342; Conservative 244; Mismatches 537; Indels 259; Gaps 47;

QY 7 VKVCVRVRLPQIRBQGOANLQWAKGNNTISQVVG--TKSFNFDREVFNSHSTSOIYQEA 65
DB 7 VKVAVRIRPLCKKALNHQVCVAVIRNSQOVLIGRVRVTFDFVFGKNTODEVYNTCI 66
QY 66 VPIIRSLAQYNGTIFAYGQTSKTYTMG-----TPNSLGIIPQAIQVFKIIOEIP 119
DB 67 KPIVLSLIEGNAVFAVYAGQTSKTYTIGGHIASVVEGOKIIPRAIOEIFQISIEHP 126
QY 120 NREPLRVSYMEIYNETVFKDLCDRRKPKLETFEDENRNVYVADLTELVMVPEHYIQW 179
DB 127 SIDNVAVSILEYKELRDLLELETSMKDLHREDKGTIVYGAKECHVESAGEVMSL 186
QY 180 IKKEKNRHYGETKMDHSSRSHTIFPMIYESRDRDPTNSNODGA--VWVSHLNL 234
DB 187 LEMGNARHGTQOMNEHSSRSHTIFISIQYHKNNEAAE--DGSWYSPRHIVSKFHF 243
QY 235 VDLASGRARQTAEGVRLKEGCINSLFLIQVYIKLSD--GQAGCFYIRRSKLTREL 293
DB 244 VDLASGRVYKTGTGTFERFESIQINSGLLAINVSALDPPRKSHPYRAKTYTRL 303
QY 294 QNSLGNAKVVIITCTIPVS--FDETSTLOFASHTAKVANTPHVNEVLDEALLKRYRK 351
DB 304 KDSLGGSAKTYMTITCVSPSSNDESNLSIKYANRANRINKTYVNFSPSSDR--IDEMEF 362
QY 352 EILDLKOLE-----NLESSETKQAMAKKEHTOLLAETIKOLHKEREDRIWH 399
DB 363 EIKLIRALQSOQAGVSOQTQINREGSPDYNRIHLSLEQVLAQEGEXXXXXXXX 422
QY 400 LTNIVASSQESQ-----QDQVYKRRRT-----WAPKIONSLH 435
DB 423 XXXXXXXXXXXXXXXXXXXXXXXXGPDARIPERRRYTVPFDTLGHYIYIPSR--ODSRK 481
QY 436 ASGVSDFDMLSRLPGNFSKKAKEFSMPPEIDSVCTEFSDFDALSMDNSGIDAENW 495
DB 482 VHSPPMYSLDRIFAGRRRSQML--LGHIEQDKVLHCGQSDMSD--EESGQKSGT 537
QY 496 LASKVTHREKTSLHOSNIDFGQISDSVQFHD-----SKENOLQYLPKDSGM 543
DB 538 RCRSRSMIOKPDSCVSLVELSDTQDETOKSDLENEIDLKIDLOESQELNLQKLNSERL 597
QY 544 AECRKASFKEKELTSLQOLOSKEEKKELVOS-----PELKIAELE--BOLSVK 590
DB 598 TEAK-----QKRELITINIKMEDLITKELGTGNDAKSVSKQSLKVTYLEDHDAEAKVE 652
QY 591 ----ANLEWVINS--REHSINAVOVDKEVEVRKESVLDGSGYNASNDLQDSVYG 644
DB 653 LIETQKQOLELENKDLSDVAMKVKILOKERRKKMDAKLAVOYLQKKQOSKILASISION 712
QY 645 KRSSSHDECIEHRK--MLQKTYDLEETLENLKKSEENDKQSSQEDFMESITOL-- 697
DB 713 EKRANLEQSVHMKYQKQIOLQKLEENENKROJDAVAKRQOKKE-----IOLKTG 766
QY 698 CEAIMAEKANALJELALMB-----DNFQNIILENETLKEIADLESLEKLENETNEFE-- 750
DB 767 QEBGLKPKAEDLDACNLKRRKSGFSIDHLQKLDQOKWMLDEVEKVLNQRLELEAD 826

```

QY 751 ILEKTEKHEAQLIHEIGSLK-KLVENAEYNNON-----LEEDL----- 789
DB 827 LKKREIYVSKKALLQEKSHLENKLRSSQALNTDSLKISTRNLNLEQELSEKNVQLQTS 886
QY 790 --ETKTLKEOEIOELAEIKRRADNLOKKVRNFDLSVSGDSEKICEE--IFOLKQSLSD 845
DB 887 TAEKTKISEQYEV---LQKEKDOLQKRHHVDDEKLNKGRVLSPEEHVLTQLEGITIA 942
QY 846 AEAATRDQKCECSFLRSENLLEKEMEDTSNNYNOKEKASLFEK-----QLETEKSN 898
DB 943 LEAAIE-----YRNESIQNRQSLRAS-FHNLISRGANVLEKLACLSPEIRTLIR 993
QY 899 YKKMEADLOKELQSAFNEINYLNGLLAGKVPBDLLSRVLEKKVSEFSKOLE-KALEKN 957
DB 994 Y-----FNKYNVNR-----EAERKQOLYNEMKMKVLERDN 1024
QY 958 ALENVTCLSEKFLPNEVECLKNOISKASEEIMLKQEGESASIIKQETIMOQESQ 1017
DB 1025 MYRELESALDHLK-----LQC-----DRLLTLOQKEHE---QKMOILLHHFKRQ 1065
QY 1018 ILQLTDEVTHT-OSKVQOTEEOYLEKKMHND-----LFEKYLIRNK-----SEADL 1063
DB 1066 DEEGIMETFEKTYEDKIQOLEKDLFYFKTSRDHKKKKELVGEAIIROQLAPEYOEGADG 1125
QY 1064 IREMENLAGTMESEVKIADTKHELEETIRKE-----QLHEKKYFQAMQTIPIPTL 1118
DB 1126 VIKRPG--GGMISEELKMAKSPRESKMLSGREMDSSASLQTPNPQKLMEDIDELPPI 1183
QY 1119 SDSL-PPSKLYEGNSODPIEINDYHNLALATERNNINVCLETENRSLEKQYIDINTOLO 1177
DB 1184 HSLAPSPGSHMLGNE-----KTEDDNQPTKSHSLSSQIQVGVNVRH 1229
QY 1178 SIQASIEKSDQKQKODLEBEVAILLEMLKGLHLDQSISTEKLEQLENEVEYKLOT 1237
DB 1230 GTVPKLCRKELRQIS-----ALELSLRSSL--GVGISMAADSIEVSRKPD 1276
QY 1238 LQ 1239
DB 1277 LK 1278

RESULT 14
US-09-641-807A-2
Sequence 2, Application US/09641807A
Patent No. 6440731
GENERAL INFORMATION:
APPLICANT: Berand, Christophe
APPLICANT: Freedman, Richard
TITLE OF INVENTION: No. 6440731el motor proteins and methods for
TITLE OF INVENTION: their use
FILE REFERENCE: 1031
CURRENT APPLICATION NUMBER: US/09/641, 807A
CURRENT FILING DATE: 2000-08-17
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 1279
TYPE: PRT
ORGANISM: Human
FEATURE:
NAME/KEY: VARIANT
LOCATION: (409)...(446)
OTHER INFORMATION: Xaa - any amino acid
US-09-641-807A-2

Query Match 5.7%; Score 841.5; DB 4; Length 1279;
Best Local Similarity 24.7%; Pred. No. 9,1e-34;
Matches 342; Conservative 244; Mismatches 537; Indels 259; Gaps 47;

QY 7 VAVCVRRVPLDREGDQANLQMKAGNNTISOVDG-TKSFNDRVFNHSESTISOIYQELA 65
DB 7 VAVAVRIRPLCKEALHNHQCVRVPIPSQOVLIGRDRVFEDPFVFGKNSDQDEYVNCI 66

QY 66 VPIIRSAQGVNCTIFAVGQTSCKTYTMG-----TPNSGILPQAOEVFKTIQELP 119
DB 67 KPLVLSLEGVAYFAVAGITTTGGGHIASVVEGQKILPRAIQELFOSISEIP 126
QY 120 NNEFLRVSYMEIYMETVADLLCDDRKKRPLIREDFNNRVVADLTELVAWPHVPIOW 179
DB 127 SIDFNKYSYIEVYKEDLRLDLEETSMKDLIREDKGNITYVAKBEHVSAEVAWSL 186
QY 180 IKKGEKNRHYGETKMNDSHRSHTIFRMIVESRDRNDPTNSNCGA-----VMVSHNL 234
DB 187 LEMGNAARHTGTQNMHESSRSHAIPTTISICQVHKMEAAE--DGSWYSPRHIVSKFHF 243
QY 235 VDLASERSAOTGAGVRLAKEGCNINRSFLITGOYIKTKSD-GAAGGFINTYDSLTITL 293
DB 244 VDLASERVTKGNTGERKESIQINSGLAGINISALGDPRRKSSHIPYDAKITRL 303
QY 294 QNSLGNNAKTVITICTTPPS--FDETLSTIOPASTAKHVRNPHVNEVLDLALKRRYK 351
DB 304 KDSLGSATVITVCVSPSSNFDESINSLKYNANRINRKPYNFSPESDR-IDEMEF 362
QY 352 ETLDLKOLE-----NLESSETKAQAKBEHTQLAEITQIHKEREDRIWH 399
DB 363 EYKLREALQSOQAGVSQTTQINREGSPDTNRIHSLREQVQALQGEYXXXXXXX 422
QY 400 LNIYVASSQESQ-----ODQRYVKKRRVT-----WAPGKIQNSLH 435
DB 423 XXXXXXXXXXXXXXXXXXXXXXXXDPDARIPERRPYVPFPHLGHYITYSR-QDSRK 481
QY 436 ASGVSDFDMLSRLPNGFNFSKAKFSDMPSEPEIDSVCTEFSDFDALSMMQNSGIDAEMN 495
DB 482 VHTSPPMYSLDRIPAGFRRSQML-LGHIEBDKYLHQCFSDNSDD--EESGEQKSGT 537
QY 496 IASKVTHREKTLHOSMIDFGQISDSVQPHD-----SSKENOLYLPKDSGM 543
DB 538 RCRSRMWLOKPPSVQSLVELSDPTQDETOKSDENEDLKIDLOEQEOLNOLKLNSERL 597
QY 544 AECRASFEKETSLOOQLOSKKEEKKELVOS-----FELKIALE--FOLSVK 590
DB 598 TEAK-----QKRELTINIKMKEDLIKELIKGNDAKSVSKOYSLAKVTKLEHDAQAAVE 652
QY 591 ---AKNLEMYTNS--REHSINAEVQTDVEKEVVERKENSVLGDSGYMASNSDLODSSVDG 644
DB 653 ILETQKQLELENKOLSDVAMKVLQKEFERKKMDAKLRYVQLQKQDSKRLASTLSTON 712
QY 645 KRLSSSHDECIRHK-----MLEQKIVDLEEFLENNKKSSENDKOKSSSEDPMESTQL--- 697
DB 713 EKRANLEQSVDMKRYOKIQLOKRLRENEKKKQLDVAIKRQOKIKE-----IOLKTG 766
QY 698 CEAIMAEKANALEELALMR-----DNFDIILENETLREIADLRSIKENQETNEFE-- 750
DB 767 QEEGLKPRAEIDDACNLKRRKGSFSSIDLQLODKQKWLDEBEKVLNOROELLELAD 826
QY 751 ILEKTEKHEAQLIHEIGSLK-KLVENAEYNNON-----LEEDL----- 789
DB 827 LKKREIYVSKKALLQEKSHLENKLRSSQALNTDSLKISTRNLNLEQELSEKNVQLQTS 886
QY 790 --ETKTLKEOEIOELAEIKRRADNLOKKVRNFDLSVSGDSEKICEE--IFOLKQSLSD 845
DB 887 TAEKTKISEQYEV---LQKEKDOLQKRHHVDDEKLNKGRVLSPEEHVLTQLEGITIA 942
QY 846 AEAATRDQKCECSFLRSENLLEKEMEDTSNNYNOKEKASLFEK-----QLETEKSN 898
DB 943 LEAAIE-----YRNESIQNRQSLRAS-FHNLISRGANVLEKLACLSPEIRTLIR 993
QY 899 YKKMEADLOKELQSAFNEINYLNGLLAGKVPBDLLSRVLEKKVSEFSKOLE-KALEKN 957
DB 994 Y-----FNKYNVNR-----EAERKQOLYNEMKMKVLERDN 1024
QY 958 ALENVTCLSEKFLPNEVECLKNOISKASEEIMLKQEGESASIIKQETIMOQESQ 1017
DB 1025 MYRELESALDHLK-----LQC-----DRLLTLOQKEHE---QKMOILLHHFKRQ 1065

QY 1018 ILQLTDEVTHT-QSKVOOTEBOYLEMKKMHDD-----LEFKYIRNK-----SEADL 1063.
 DB 1066 DEGEIMETFEKYEDKIQLEKDLFYKKTSDHKKKKEIKVEIARROLAISEYQADG 1125
 QY 1064 LREMENLKGTMSEVYKIDTKEHELEETIROKE-----QLHEKKYFQAMOTIPIPL 1118
 DB 1126 VLKPEG--GGMISEELKMAKSPRESKLSGREREMDSASLSRTQPNPOKLMEDIPELPI 1183
 QY 1119 SDSL-PPSKLVGNSODPLEINDYHNLJALATERNNINWCLETERNSLKEQVIDLNTOLQ 1177
 DB 1184 HSSLAPSGHMLGNEN-----KTEIDNOFTSHSLSSQIOVGVNGRLH 1229
 QY 1178 SIQAOISIEKSDIQKODLEGEVYLLMEILKCHLTDLSQLEKLOLENIETVEKLTQ 1237
 DB 1230 GYTPVKLCRKEILROIS-----ALELSIRSSL--GVGISMAADSIIEVSRKRPD 1276
 QY 1238 IQ 1239
 DB 1277 LK 1278

RESULT 15

US-09-723-096-2
 : Sequence 2, Application US/09723096
 : Patent No. 6448026
 : GENERAL INFORMATION:
 : APPLICANT: Beiraud, Christophe
 : APPLICANT: Freedman, Richard
 : TITLE OF INVENTION: No. 6448026el motor proteins and methods for
 : FILE REFERENCE: 1031
 : CURRENT APPLICATION NUMBER: US/09/723, 096
 : CURRENT FILING DATE: 2000-11-27
 : PRIOR APPLICATION NUMBER: US/09/641, 807
 : PRIOR FILING DATE: 2000-08-17
 : NUMBER OF SEQ ID NOS: 4
 : SOFTWARE: FastSeq for Windows Version 4.0
 : SEQ ID NO 2
 : LENGTH: 1279
 : TYPE: PRT
 : ORGANISM: Human
 : FEATURE:
 : NAME/KEY: VARIANT
 : LOCATION: (409)-(436)
 : OTHER INFORMATION: xaa - any amino acid
 US-09-723-096-2

Query Match 5.7%; Score 841.5; DB 4; Length 1279;
 Best local Similarity 24.7%; Pred. No. 9,1e-34;
 Matches 342; Conservative 244; Mismatches 537; Indels 259; Gaps 47;

QY 7 VKVAVRVRPLQREOGQANLQWKAGNNTISQVDG-TKSFNDRVENSHESTSQIYOEIA 65
 DB 7 VKVAVRVRPLQREOGQANLQWKAGNNTISQVDG-TKSFNDRVENSHESTSQIYOEIA 65
 QY 66 VPIIRSLQNGTITFAVYGGTSSGKTYTMG-----TPNSLGIIPQALQVKKIIOEIP 119
 DB 67 KPLVLSLIEGYNATVAFVYGGTSGKTYTIGGHIASVYEGQKGIIPALQEIOTISIEHP 126
 QY 120 NREFLLVSYMEIYNETVTKLDCDRRKKPLREDFENRNVVYADLTELVAPEHYIQW 179
 DB 127 SIDFNVSYSIEYKEDLRLELETSMKDLHIREDEKQNTVYGAKECHVESAGEVMSL 186
 QY 180 IKKGKRRHNGETKRNDSRSHITITPMYVESDRNDPTNSGCGA-----VMVSHLNT 234
 DB 187 LEMGNARHGTQOMNEHSSRSHAFITISICQYHKKMEAAE--DGSWSPRHIVSKFHP 243
 QY 235 VDLAIGERASQTAGEGVRLKEGGINRSLFIILGOVTKKSLD-CQAGGFINRSKATRLI 293
 DB 244 VDLAIGERASQTAGEGVRLKEGGINRSLFIILGOVTKKSLD-CQAGGFINRSKATRLI 293
 QY 294 QNSLGNNAKTVIICITTPVS--FDETLSTLQFASAKHVNTPHVAEVLDEALLKRYRK 351
 DB 294 QNSLGNNAKTVIICITTPVS--FDETLSTLQFASAKHVNTPHVAEVLDEALLKRYRK 351

DB 304 KDSLGSAKTYMTCVSPSSNPFDESLNLSKYANRANINRKPVTNPSPESDR-IDEMEF 362
 QY 352 EILLLKQLE-----NLSSESTKQAMAEHEHTQLAETIKQHKEREDRTW 399
 DB 363 EIKLIRALQSQAQVSGQTOINREGSPDTRIRHLSIEEQVAQLQEGEXXXXXXXXXXXXX 422
 QY 400 LFNIVASQSQSQ-----ODORYKRRRT-----NAPKIQNSLH 435
 DB 423 XXXXXXXXXXXXXXXXXXXXXXXDPAKIPERRPYTFPHLGHYIYPSR-ODSRK 481
 QY 436 ASGVSDPMLSRIPGNSKAKESDMPSPEIDSVCTEFSDFDALSMDSNGIDAENW 495
 DB 482 VHTSPMYSLDRIPAGFRFSQML-LGHIEQKVLHLCQSDMSD--EESGQKSGT 537
 QY 496 LASKVTRERTSLHQSMDIDEGQISDSYQFD-----SKENOLQYLPKDSGM 543
 DB 538 RCRSRWIOKRPDVSGLVELSDTQDETQKSDLENEDKIDLOESQELNQLKNSERLI 597
 QY 544 AECRASFEKEITSIAQOLSKKEEKELVOS-----PELKIAELE--EQLSVK 590
 DB 598 TEAK-----QKRELTINIKKEDLIKELIKTGNDAKSVSKYSLVTKLEHDAQAYE 652
 QY 591 ---AKNLEAVTNS--REHSINAEVQTVKEVYRREMSYLAGSGYNASNSDLQSSVDG 644
 DB 653 LIETQOLQLELNKDLSDVAMKVKLOKEFRKKMDAKLRVQLQKQOQSKLASLQIN 712
 QY 645 KRLSSHDECIEHRK-----MLBOKIYDLFEITNLKKSNDKQKSSODFMESTOL--- 697
 DB 713 EKRNLEQSVSDHMKYOKIOLQKRLRENERKROLAVIKRQOKKE-----IOLKTG 766
 QY 698 CEAIMAEKANALEELALMR-----DNFDNITLENETLKEIADLERSLKENOTNEFE-- 750
 DB 767 QEGGLKPKRADLDACNLKRRKSGFSIDHOKLDEQKNMIDEVEYKVLNQRQLELEND 826
 QY 751 ILKEKQKEHEAQLIHEIGSLK-KLYENAEVNTQN-----LEEDL----- 789
 DB 827 LKRRKIVSKREALQEKSHLENKKLRSSQALWTDSLKISTRNLLEQELSEKNVOLQTS 886
 QY 790 --ETKTLKEOEITQLELTKRRADNLQKVRNPLDSVSGDSKLEE--IFOLKQSLSD 845
 DB 887 TAEKTKRISQVAV-----LOKEKQOLQKRRHDVDEKLNKGRVLSPEEHVLTQLEGGIA 942
 QY 846 AEAVTDAQKEGSLRSENLLEKEMEDTSNMYNQEKASLEEK-----QLETEKSN 898
 DB 943 LEAAIE-----YRNESIONRQKSLRAS-FHNLISRGANVLEKLACLSPEIRTIIFR 993
 QY 899 YKKGMEADLOKELQSAFNEINYLGLAGVPRDLRSVLEKYSFSGQLE-KALEEN 957
 DB 994 Y-----FNKYVNLN-----BAERKQOLYNEMKAKVLEEDN 1024
 QY 958 ALENVYCLSEYKFLPNEVECLKNQISKASEIMLKQEGEHSASLISKOETIMOQSQ 1017
 DB 1025 MYRELSALDHUK-----LQC-----DBRLTLQCKEHE--QKQMLLHHPKQ 1065
 QY 1018 ILQLTDEVTHT-QSKVOOTEBOYLEMKKMHDD-----LEFKYIRNK-----SEADL 1063
 DB 1066 DEGEIMETFEKYEDKIQLEKDLFYKKTSDHKKKKEIKVEIARROLAISEYQADG 1125
 QY 1064 LREMENLKGTMSEVYKIDTKEHELEETIROKE-----QLHEKKYFQAMOTIPIPL 1118
 DB 1126 VLKPEG--GGMISEELKMAKSPRESKLSGREREMDSASLSRTQPNPOKLMEDIPELPI 1183
 QY 1119 SDSL-PPSKLVGNSODPLEINDYHNLJALATERNNINWCLETERNSLKEQVIDLNTOLQ 1177
 DB 1184 HSSLAPSGHMLGNEN-----KTEIDNOFTSHSLSSQIOVGVNGRLH 1229
 QY 1178 SIQAOISIEKSDIQKODLEGEVYLLMEILKCHLTDLSQLEKLOLENIETVEKLTQ 1237
 DB 1230 GYTPVKLCRKEILROIS-----ALELSIRSSL--GVGISMAADSIIEVSRKRPD 1276
 QY 1238 IQ 1239
 DB 1277 LK 1278

Sun May 25 14:36:17 2003

us-09-150-867-1.rai

Page 19

Search completed: May 23, 2003, 14:51:52
Job time : 77 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2003, 14:40:15 : Search time 282 Seconds
(without alignments)
1038.851 Million cell updates/sec

Title: US-09-150-867-1

Perfect score: 14769
Sequence: 1 MSEGDAVKVCVRRPLQRE.....QAENMWYEAKKETAPCKTS 2954

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 375593 seqs, 99172665 residues

Total number of hits satisfying chosen parameters: 375593

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
5: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
6: /cgn2_6/ptodata/2/pubppaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep:*
9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep:*
11: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1015	6.9	3899	9	US-10-171-311-4
2	1015	6.9	3917	9	US-10-171-311-8
3	1013	6.9	3907	9	US-10-171-311-2
4	1013	6.9	3925	9	US-10-171-311-6
5	993	6.7	2383	9	US-10-082-830-260
6	821	5.6	2334	9	US-09-820-843A-113
7	811	5.5	2310	9	US-09-991-496-120
8	811	5.5	2310	9	US-09-820-843A-114
9	811	5.5	2310	10	US-09-874-923-110
10	794	5.4	2139	9	US-10-023-219-4
11	794	5.4	2139	10	US-09-727-384-6
12	761	5.2	1805	9	US-09-820-843A-73
13	720.5	4.9	2835	10	US-09-885-535-4
14	719.5	4.9	1979	9	US-09-927-597-4
15	718	4.9	6281	10	US-09-815-242-12996
16	717.5	4.9	1972	9	US-10-171-311-162
17	708.5	4.8	1945	9	US-09-927-597-2
18	706.5	4.8	1938	9	US-10-171-311-164
19	691.5	4.7	3158	10	US-09-815-242-12611

20	651	4.4	2025	10	US-09-815-242-5703	Sequence 5703, Ap
21	651	4.4	2437	10	US-09-815-242-5834	Sequence 5834, Ap
22	649	4.4	1103	10	US-09-847-874A-1	Sequence 1, Appl1
23	643.5	4.4	1786	9	US-09-742-096-3	Sequence 3, Appl1
24	624	4.2	2125	10	US-09-919-172-29	Sequence 29, Appl1
25	567	3.8	864	10	US-09-883-096-2	Sequence 2, Appl1
26	555	3.8	383	10	US-09-883-096-5	Sequence 5, Appl1
27	544	3.7	460	10	US-09-925-300-1228	Sequence 1228, Ap
28	527	3.6	338	10	US-09-883-096-4	Sequence 4, Appl1
29	526.5	3.6	1993	9	US-10-098-979-2	Sequence 2, Appl1
30	525.5	3.6	2285	10	US-09-932-183A-2	Sequence 2, Appl1
31	522	3.5	868	9	US-09-884-001-19	Sequence 19, Appl1
32	516	3.5	2055	9	US-10-017-216-4	Sequence 4, Appl1
33	514	3.5	5795	10	US-09-815-242-12610	Sequence 12610, A
34	500.5	3.4	2053	9	US-10-017-216-2	Sequence 2, Appl1
35	496.5	3.4	1591	9	US-10-197-666A-92	Sequence 92, Appl1
36	495.5	3.4	1958	12	US-10-028-946-4	Sequence 4, Appl1
37	495.5	3.4	2054	12	US-10-028-946-2	Sequence 2, Appl1
38	484.5	3.3	1203	9	US-10-097-340-43	Sequence 43, Appl1
39	481	3.3	1641	9	US-10-017-216-5	Sequence 5, Appl1
40	480.5	3.3	1597	9	US-10-017-216-6	Sequence 6, Appl1
41	477.5	3.2	2478	10	US-09-815-242-5816	Sequence 5816, Ap
42	477.5	3.2	2478	10	US-09-815-242-12967	Sequence 12967, A
43	474.5	3.2	890	9	US-10-171-311-196	Sequence 196, App
44	455.5	3.1	1368	9	US-10-159-151-2	Sequence 2, Appl1
45	447.5	3.0	3788	9	US-09-952-267-76	Sequence 76, Appl1

ALIGNMENTS

RESULT 1
US-10-171-311-4
Sequence 4, Application US/10171311
Publication No. US20030087270A1
GENERAL INFORMATION:
APPLICANT: Schlegel, Robert
APPLICANT: Chen, Yan
APPLICANT: Zhao, Xumei
APPLICANT: Monahan, John
APPLICANT: Kamatkar, Shubhangt
APPLICANT: Glatz, Karen
APPLICANT: Gannavarapu, Manjula
APPLICANT: Hoersch, Sebastian
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
TITLE OF INVENTION: OF CERVICAL CANCER
FILE REFERENCE: MRI-035
CURRENT APPLICATION NUMBER: US/10/171,311
CURRENT FILING DATE: 2002-06-12
PRIOR APPLICATION NUMBER: US 60/298,159
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/298,155
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/335,936
PRIOR FILING DATE: 2001-11-14
NUMBER OF SEQ ID NOS: 238
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 3899
TYPE: PRT
ORGANISM: Homo sapiens
US-10-171-311-4
Query Match 6.9%; Score 1015; DB 9; Length 3899;
Best Local Similarity 19.6%; Pred. No. 6.2e-31;
Matches 700; Conservative 630; Mismatches 1164; Indels 1072; Gaps 148;
QY 181 KKGGKNHYETKNDSSSHRTFRIVSRNDPTNSNCDGAMVSHLNLVDLAGS 240
DB 34 KKQKKR--KTSSSKHDVSAHDLNIDGOCNMYTINSSQVRVESTVPIPESTIMRTLHSG 90
QY 241 ERASQTAEGVRLLK-----EGCN--INRSLFLIGGVIRKLLSDGQAGFTINRDSKLT 291

```

Db 91 EITSEOGFSEVLESEISTTADCCSEYNGCSFYM-----RTGFPIN 132
QY 292 ILQNSLGNATVITICITTPVSPELTISTLOFASTAHVNTPVHNEVLDDEALLKRYK 351
Db 133 LIREBERG-----YDYS-----EGADSPHLMMESELAGKOH-- 169
QY 352 EILDIKOLENLESSSETKA-----QAMAKEH---TOLLAIKOLHEREDRIMHLT 401
Db 170 EIELNLEEMRYTGTGEGLOJQOEPEALIKORDGIITQITANLQOAREKDETFMEFL 229
QY 402 NIYVASSOESQODRVKRRVTPAPKIONSLHAGVSD-----FDMLSR 447
Db 230 EL-----TEOSQKIQIOFOOLQASSETLRNSTHSTAAIDLQAKQIILTHQOOLEBOD 282
QY 448 LPGNFSK-----KAFSDMSPFEIDSCVCEFDGFDALMMSNGDAE 493
Db 283 LREYQKKKDEFTMOJFLOEKIKYEMEDCKVENSKEEIOEKETIIEIMTKIIEE 342
QY 494 ---WNLASKVTHREKT---SLHQSMDFGQ-----ISDSVOFHDSKENOLQYLPKDS 540
Db 343 KTLLELDKDLTTADKLGEIOEQIVQKNQIKNMKLELTNSKQKERSSEIKQIM----- 398
QY 541 GDMAEGRK-----ASPEKETS--LOOQLOSK--BEKKELVQSELTAELEBOL--SVKA 591
Db 399 GYVEBELQKRNHKDSQFETDIVQREDETOKRLQOLRAELDEMGQOIVQMKOELIRHOMA 458
QY 592 KNEVYNTSR---EHSINAEVQDVEKEVYVRKESYVLDGSGYNASNDLODSVQDKRS 648
Db 459 QMEEMKTRHKGEMENALRSYSNITVNDQIK-----LMNVAINELNIKQDITNSQKELK 513
QY 649 SMDCEIEHKKMLQKTVLDEEFIEINLKKSEMDKQSSBODM-----ESIOLCALIMA 703
Db 514 EEBGLILEECAJQROLELVEELSFRQJQIARQITAEQESKLBAHNSISTVEDLKA 573
QY 704 EKANALE---ELAL-----MRDNFDNIIENETLKREIADLER----- 738
Db 574 EIVASASRELELKHAETVANKIKLEMLEKKNVLDMAESQ--AELEKRIOTLLS 632
QY 739 ---SLK-----NOETNEFEI--EKETOKEHE 761
Db 633 HEELSKLEDELEIHRINIEKLKDNIGIHYKQOIGLQNMESQKJETMQFERNLITRQ 692
QY 762 AOLIHGISTKLVENAEMYNONLEE-----DLETKTKLKE-----QEIOL 803
Db 693 NOILIEISKLQDOOS--LYNSKSEMTQJQINLOKEIETLQOEKEKQETLDEQVEIOL 750
QY 804 -----AELRRADNLQKVRNFDLSVS--MGDESEKLCEEIFOLKOSLSDAEAV----- 849
Db 751 KTELLEKQMKERENDQOEKFAOLEAENSLTKDEKTLDEMILHHPVQSEERLIFIDSTK 810
QY 850 ---TRDA--QKESFLASEN-----LEKKEKEDTSNNYNOKERKAASLFQOLETKSN 898
Db 811 SKSKDSVWKEKEIETILEENEDLKQOOCIQINEIEKORNTFSFAKNEFVNOJOLEBYAC 870
QY 899 YKMEADLO---KELQ-----SAFNEINYLN----- 921
Db 871 LKVVMDLLEDSKNQOLEKSKKLALNEELHJQRIINPTVKKMSYFDEDKTFFVAETLEM 930
QY 922 GILLAGKVPDDL-----SVALEKKVASESKOLEKALEKKNALNEVTOLESEYK-- 970
Db 931 GEVVEKQDTTELKLEVTREKRELTQSRLSDSEOLKQNGEISPLNEVYVSLKQEKQOY 990
QY 971 -FLPREVELEKYN-----QISKASEIMLIKQEG--EHSASITSK----- 1006
Db 991 SLRCELEETIINHNAENVOCDTQVSSILDDGVYMTSGAGSVSKYKSGEESKIMV 1050
QY 1007 -----QEIINOJQSEQ---ILQTFDEVYTHQS--KYQOTEOBYE-----MKKHNDLF 1050
Db 1051 EKVAFEMNTVGEESKQOLILDHIPSTYKESLSLAPQPSKNDKQLEKLVNLSKSEOND 1109
QY 1051 EKYTNKSEAE-----DLRRE--MENLKT-----MESVEVKIADTKH 1086

```

```

Db 1110 -----RQMEQRICLSLVSTHVDQVREYNEKKAALCSLEELIFAQOEKIKELQKH 1165
QY 1087 ELE-ETIR-----DKBOLLH-----EKYFFQAMOTIPITPLPSLSPSKLY 1128
Db 1166 QLEIQTMKTOETQDEDEKRPHLILGLKQKAVSECSYFLQTLCSV-----LGEYTPALKC 1220
QY 1129 EGNSDPIELINDY-----HNIALATEKNNTMVCLEF----- 1160
Db 1221 EVNAEDKENSQDITSENEDEPELODYREVOQDPOENHNTLNKVEYNNLLVLIQRLSKI 1280
QY 1161 -----ERNSLKEOV-----IDLNTLOSLQASIEKSDLOPK 1193
Db 1281 WGOQTQGMKLEGEENLPKEETEFPLSHSQMNLBIDVNHHSKSLSLQDLEKTLIEBY 1340
QY 1194 QDLE-----EGEYKLLIEMLKCHL-----TDSQSLT 1221
Db 1341 QELISLISLQOOLKTEBQNEVBAEHICQKRLQAVSESTVPSLPVDSVITRESDAORTM 1400
QY 1222 -----EKLQENLEVTLEQJLOEBEMKNIT-----ERNE 1251
Db 1401 YGSCYKKNIDGTIERSGEYKKEETINYKLEKOQOELEEVAKVYSMSIAFAQOTE 1460
QY 1252 L-----QTFEDLKAHDSLKQDLSENIEOSIETODELRAAQ 1288
Db 1461 LSRISGKENTASSKQAHAVCOEOHYFENMKLSODQIGQFETIVD--VKFEKFPKLS 1518
QY 1289 FELREO--KQLYDSFRQ-----OLDCSVGIS-----SPNHDAVANO---EKSISLGVNS-- 1333
Db 1519 KELGEHKGITILSNSDPHDIPESKDCVLTISEMFSKDTFTVPKOSIHDEISVSSMDSR 1578
QY 1334 ---LOSEMLGRDELQJSCALVSELELLR--AHVKSVEGENLEETKRLNGLEKEILKS 1389
Db 1579 QJLMEQOELEDRQELVROVQEHQOATELLRQAHNRQHE-----RQ 1619
QY 1390 ESEVYKSLMLEKEDNKTKQEAEEYSKENQFLEEVFGSQKLYVIEYLAQKAA 1449
Db 1620 REDQ-----EQLQOEIKRLNQLAQRSSIDENIVSE--REYVLELALQKOLSLAG 1670
QY 1450 EERL--EIKDRYFELVQNTANTNIVEGLKLETPLOADHEDSDIRSEEMETKVGELKER 1507
Db 1671 REKLCCELRNSS---TQTONENQGEVE--EOTFEKELDRKREVPPEILSNE--- 1720
QY 1508 NOYLLERLOEKELESNKLEILOKEMETSVILKDD-----LOQKLESISENIILKE 1559
Db 1721 -RYALQKAN-----NRILKILLEVYKTTAAVEETIGRHVIGIIDRSSKSSOASLWRS 1773
QY 1560 NIDTTLK---HHSQOALQOLK-----TOQEOJQAKNL---ATAAS 1593
Db 1774 EAEASVKGCVHEHRTVYDSEISPTSGSDMRNDINMMSKYTERGETLSQRLVNSGFRGT 1833
QY 1594 DNCPIQOE-----KETSAQVHPLEKILLLTEELHOKTNEOKL 1633
Db 1834 EIDPENEBELMINISSRLQAAVEKLLEAISETSQLEHAKVYQTELMESPFRQKEAFESL 1893
QY 1634 -----LHEKNEL--BOQAVEKCEVYEHLMKMSIESKSSLESLOHEKHD---TEQOL 1679
Db 1894 KQOELEIRLHEESAPARQALAVELS--KAEVYIDQYAEKTLFEQIOIEKTDIIDRLQOEL 1952
QY 1680 LALQOQOVVTOQEKKELOQTEHNTLAEVYDHLKENIEIGLNFKNNAQOKTKTQEOCLLENK 1739
Db 1953 LCASNRLOELABEQOQIOEBELLSROKEMAK--AEAG-----PYEQOLLOETE 1999
QY 1740 ELEQSQRHLOCEIEBELKSLKDKESALETLESBQKYVINLQOEMENYMLEMEELKNSQRT 1799
Db 2000 KIMKEKLEVOQOAKVYDDLOQYKALE-----IDVEQYSR--FILEQEKNELEM 2049
QY 1800 VIARQDLODDLRSEVEMSIETQDLDKQAEALQOQKQVQELTSQISVL----- 1849
Db 2050 DLROQONALEKQELKMKRFTLDEQAIDREHBRDYQOQ--EIQKLEOQKLVPRPROPLISEHQ 2107
QY 1850 QEKISLLENOMLVAVVUKETILSERDLONSKQHLFSEITILSLKKEKFAI----- 1902
Db 2108 TREYQOLANHLKEXTDKCSEILLSKQOLOKQJQERNEIEKLEFRVRELEQOALLVEDRKH 2167

```

```

QY 1903 -----EOAKKADAKRTIDITEK--ISNIEOILLQATNKEPIYERESL--IQCKE 1952
Db 2168 FGAVEAKPELSVEQLOAERAPADIKEREKTEITNEEOLEQREBELNKEEVOQLOHMOLEI 2227
QY 1953 QALANTEHELETTLSKDLALCKME-----QOEDEANKVIALTEKMSLSE 1998
Db 2228 OKKSTTRLOLEBQENKLFNDMEKLGAIKESDAMSTQODHVLFGFPAQIIOKEVEID 2287
QY 1999 QINENTTTLEBGEKETEYLQRPSSQOOSOMELNBSLTKDLOLE-----E 2047
Db 2288 QLNQOAVYKLO-----QOLKITTDNKVIEKNEMLIDTLETOELCMLSDQECYKN 2336
QY 2048 AEKESATNEIKNLAKISLSEELIIONASILNEAVSERENLHNSQOVLSE---LEO- 2103
Db 2337 REEIEOLNEVIEKLOELANIGOKTSMNAHSLSE---EADSLKHOLDVYIAEKLALEOQ 2393
QY 2104 -----LSLTLKSRDHAFQ-----SKREKDEAVNKIASLAE-----IKILTK 2141
Db 2394 VETANEMETPMKNVLEKTFNFKMQLJQELPSLKRER--ESVEKIOSIPENSYNVALDHLK 2452
QY 2142 EMDEF-----RDSKESLQEOSSHUSE-----LCYTKTELQMLKQOKEDINKLAEKVK 2190
Db 2453 DKPELEVLTEDALKSLLENQTYEKFSPENGKSGIINLETRLQLESTVSAKDELTOCYK 2512
QY 2191 EV-----DELLQ-----HLSLKEQ-----LDQIOMELANE----- 2216
Db 2513 QIKMODOGOFETEMLOKKTIVNLOKIVEEKVAALVSOIOLEAVOEAKFCQDNQITSSSE 2572
QY 2217 -----KLRYVELCEKMDIMEKEISVLRLOQNEP-----QOEEDVAERMDILES 2260
Db 2573 PERTINIONLQNRDELQSDISATLRISELESQVEMHSLIIEKQVETAEK--NVLEK 2631
QY 2261 RMO--EIOELMEKISAVYSEQHTLSSLSSELOKETEPAHKHCMNLIKESLSTLSRSGS 2318
Db 2632 EKKLELEQKLEEGNEKKOREKREKRSPODVEVLTETTELFEF-----SNESSGFNELEA 2685
QY 2319 LQPEHVAKLNTQLOTLNKKFVYVRYTAAYKEDH--SLIKDE--KDLAEOK-----RH 2367
Db 2686 LRASVATKRAELASYKKAELQLOELLYKETNMTSLQKDSQVHDHLAEKESLIEKE 2745
QY 2368 DELRLQ--LOCL--EOHGRKMSDASEE---LKCEIEFLNELLFKRANTIQSVODFS 2419
Db 2746 DEFVQSKKACMEPRLPIKLSKISASQTDGTLKISSNQPOLVKNAGIQIOLNSPCS 2805
QY 2420 --EYQVPLNOVGSTLOELEHKKGFQMLBEFGDLHDAKLSGMOENRIST----- 2473
Db 2806 SEEVTELIISO-----FTEKIEKMOELHA-----ABILLMESHHISETEFLK 2846
QY 2474 -----IOLLTKR--LKAVVQ-----SKI-----OREIT-----VY 2496
Db 2847 REHVAVOVLKKECGITKAVIQCRLKSGSSIPELASDAVQTEILCSSDSGSDMGOCIT 2906
QY 2497 LNO--FEARLOEKKRON-----KELMRMEHGPSA-----SYM 2538
Db 2907 LTHSGQDIASEGGESESATDSFPKKIKGLRAVHNEGVOVLSLTFESPJDSGEDHSIQ 2966
QY 2529 E-----EENARLLGILKTYODESKLQSRKMLNENLNVKDDAM-----KGEVAYI 2576
Db 2967 QVSEPMLEERKAYINTISLKLITKMO--LQRAEYVDSQSSHESFSDMGDELILA 3031
QY 2577 LOKL-----SRNAEALNMOVKLTFKQONLDAAM----- 2609
Db 3022 LQOVLFEERSVLLAAPTETLTALCTDAVGLNCLQLEORIQOGEVYQAMECLOKADRRS 3081
QY 2610 --EENIOMKAVAGAVPYKEEID-----NLTK-----VYKIEKIKYSKAT 2650
Db 3082 LLSIQLALHOMNQRKLTLRQESSEKPSQELLEYNIQKOSQMLQWGVLSMK--DRAT 3140
QY 2651 DQELAYLKSCLDEKEELRRIKELRRAPQADNDTVCVPKDYQASITFPYTCGGSGSIGVQ 2710
Db 3141 E-----LQEBLSSEKMAVVAELKSELQTKLETTLLAKOHRHLK----- 3179

```

```

QY 2711 STAMVLQSEKALERELSHYKKKKYHHLSPRTSSSEDR-----KTKKASDAHSSHT 2762
Db 3180 -----ELEAFRLKVKDXTDEVHLLNDTLASQKSRLOALALEKRAK-----L 3223
QY 2763 GSHRSGSPKTYETRYNHG--PVPPERSEMPSLHSGPKKSSSTKRVASPNRSEIYSLVMS 2821
Db 3224 GRSEBRKEELEDLKFSLESQKORNLQNLLEQOKOLNBSQOKIESQRLYDAQLSSE 3283
QY 2822 PKGTGMKHILSPSKVGLHKRRLASPNRSEMPYQHVHSPKGTGLKHNLTSTLFDNLSSP 2881
Db 3284 QGRNLEQVLLSESEKVRRLRENSSTLDRERELHAQLOSSDGTGQSRPPLPSBLLKELQK 3343
QY 2882 C--KOQKVOENLN--SPKGLFDVSKK 2904
Db 3344 LEKHSRLVELNTEKTKKLDISQTR 3369

RESULT 2
US-10-171-311-8
; Sequence 8, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangl
; APPLICANT: Gialt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoerish, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171, 311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 3917
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-171-311-8

Query Match 6.9%; Score 1015; DB 9; Length 3917;
Best Local Similarity 19.6%; Pred. No. 6.2e-31;
Matches 700; Conservative 630; Mismatches 1164; Indels 1072; Gaps 148;

QY 181 KKGKRNHYGTCKKNDHSSRSHTTFRMIVESRDRNDPTNSGCGAVVSHLNLVLAGS 240
Db 34 KKQKKRR--KTSSKHDVSAHDLNDIQSCNEMYINSQSVSTYIPRESTIRTLHSG 90
QY 241 ERASQTAEGVRLK-----EGCN--INNSFLIGOVYIKKLSGQAGGFTNYDSKLR 291
Db 91 EITSHQGFVSVELSEISTADDCSSSEVNGCSFYV-----RTGKPTN 132
QY 292 ILQNSLQGNATVYICITTPVSFDETSLQFASTAKHVNRTPVHNVYDEDEALLKRYRK 351
Db 133 LIREERFG-----VDSYS-----EGQADSPHLEMESELAGKH-- 169
QY 352 IILDLKOLENLESSSETKA-----QAMAKEEH--TOLLAEIKOLHKREDRIHMLT 401
Db 170 EIEELNLEEMRYTYGEGLOQLOFEFPAIKQDGIITQTLANLQCARREKDTMREFL 229
QY 402 NIVVAOSSQODQVRKKRRVYAPGKIONSLSASGVS-----FDMLSR 447
Db 230 EL-----TEOSOKLQIOFOLOASSETLRNSTHSTRAADLLQAKOOILTHQOOLEEODH 282

```

QY 448 LPGNFSK-----KAFSDMPSPFIDSDVCTEFSDFDALSMDSDNGIDAE 493
 Db 263 LLEDYQKKKEDFTMOISLQEKIKYEMHODKVENSKKEILOEKTITIEELNTITIEE 342
 QY 494 ---WNLASKVTHREKT--SLHOSMIDFGO-----ISDVQFHDSSKENDOLYLPKDS 540
 Db 343 KKTIELKOKRTIADKLIGELQBOIVQKNQOETIKMMKLELTNSKOKROSESEIKOLM---- 398
 QY 541 GDMACSR-----ASFEKEITS--LOQOLSK--EEKKKLVSOFELKIMLEBOL--SYKA 591
 Db 399 GIVELOQRNKHDSQFEDYIQRMEQOTRKLEOLRALDEMYGGQIYQMOOEILRQMA 458
 QY 592 KMLEVNTSR--EHSINAEVQTVKEVVRKENSVLDSGYNASNDLOPSSVDGRKLS 648
 Db 459 QMEKRTKRGEMENALNSYSNITVNEQIK-----LNVAINELNTIQTNSOKETKL 513
 QY 649 SSHDCTIHRKMLBQKIYDLEEFLENLKKSENDKOKSEDDFM-----ESIQLCEALMA 703
 Db 514 BELGITLEKCALOQLELDLVEELSFSRQIQRARQITABOESKLINEAKHKSISTVEDIKA 573
 QY 704 EKANALE---EIAL-----MRDNFIITILENTEKREIADLER----- 738
 Db 574 EIVASSESRKLELKEHAEVYNTYKIMLEKERNAYIDRAESQF--ALERLTQILFS 632
 QY 739 ---SLKE-----NOETNEFEIL--EKETOKEHE 761
 Db 633 HEEELSKLEDELEIHRINIEKLDNLGIDGIONEMSOIKETMOPEKDNILTKQ 692
 QY 762 AOLIHEIGSLKLVENAMYNONLE-----DLETYKTKIKE-----OEOL 803
 Db 693 NQILLETISKLDLOS--LVNSKSEMTLOJNELOKEIILROEKEKGTIODEYLOEL 750
 QY 804 ---ALRRRADNLQKRVNFDLSV--MGDSEKICEEITQOLKOSLDAV----- 849
 Db 751 KTELLEKQKKEKENDLOEKFQALEAENSILKDEKTKLEDMKIKITPYSOERLITLDSIK 810
 QY 850 ---TRDA--QKRECSFLSEN-----LEIKEMEDTSMNTNOKKAASLEFEKOLETKSN 898
 Db 811 SKSDQSWMEKIEITLIEENEDLKQOOCIOLENEEIKORNTESFAEKNEVYVOELOEYAC 870
 QY 899 YKKMADLO---KELO---SAPNEINYLN----- 921
 Db 871 LTKVQDLEDSKKNQOLEYKSKLALNEBILHOKINPTVYKMSVYDEKTIYAETLEM 930
 QY 922 GILAGKPRDIL-----SRVLEKKVSEFSKOLEKALEKKNALENEVCLSEYK-- 970
 Db 931 GEVEXKDTTELMEKLEVTYKREKLELSQRLDSEOLKQHGELISFLNEEYKSLKOEKEQV 990
 QY 971 -FLPNEVECLKN-----QISKASEEITMLKQEG--EHSATITSK----- 1006
 Db 991 SLRRELEIILNHRANVOSCTOVSLLDGYVMTSRAGESSVKVNSFGESEKIMV 1050
 QY 1007 -----OETIMQOSEQ--TLOLTBEVHTQOS--KVQOTEEYOILE-----MKMHDDLF 1050
 Db 1051 EDKVSFENMTYVGEESKQOELILDLHPSVTKESSLRATOPSENDKLOKELNVLSQENDL- 1109
 QY 1051 EKYIRKNSGEA-----DLRE--MENLKG-----MESEVYKIADTKH 1086
 Db 1110 ---KLQMEKQORICLSLVSTHVQVREYEMENKRNALCSLKEELLEFAOEKIKELQKIH 1165
 QY 1087 ELE-ETIR---DKBQLH-----EKYFFQAMOTIFPTPLSDLSPLPSKIV 1128
 Db 1166 QLELOQNMKTQETDEGKPHLLGLKQKAVSECSYFLOTICSV-----LGEYTPALKC 1220
 QY 1129 EGSQOPRIELNDY-----HNTIALATERNNIMVCLT----- 1160
 Db 1221 EVNAEDKENSQDIYSENEDPELODYREYVODFOENNMHTILNKYTEYKRLIVLQTRLSKI 1280
 QY 1161 -----ERNLSKEOV-----IDNTQOQSIOQASIEKSDLOKAK 1193
 Db 1281 WGOOTQMKLEFGEENLPEETEFILSIHOSMTNLDEIDVNHNSKLSLODLEKTKLEOV 1340

QY 1194 ODLE-----EGEVKILLEMELKGLH-----TDSOLST 1221
 Db 1341 QELSLSSLSLOQLKETEONTEAETHLOKRLQAVSESTVPPSLPVDVYTSDAQRTM 1400
 QY 1222 -----EKLOENLEVETKLOTLOEKKNTTI-----ERNE 1251
 Db 1401 YPGSCVKNKIDGTIEFSGEGVKEETNIVKLEQYQBOLEEVAKYVMSIAFAQOTE 1460
 QY 1252 L-----QTNFEDLKAHSDLSKODLSENTIBQSIETODELRAAQ 1288
 Db 1461 LSRISGKEWNTASSKQAAVQOQEHFYENMKLSODQIGQTEFETV--VKFKEEPKPLS 1518
 QY 1289 EELREQ--KOIYVDSFRQ-----QILDCSVGIS-----SPNHAVANO---EVSIGEVNS-- 1333
 Db 1519 KELGEHKEKILLSDPHDIPESKDCVLTISEEMFSKDKTFYIROSITHIEISVMSDADR 1578
 QY 1334 ---LOSEMLRGERDELOTSKALVSELELIR--AHVKSVEGENLEITKYLKLEKEILGKS 1389
 Db 1579 QLMINEBQLEDMRQELVROYQEHQOATELLRQAHMRQME-----RQ 1619
 QY 1390 EESVVLKSMLENKEDNNKIKEQAEESYKSKENQPSLEEVYSGQKYVDEIYVLAQOLKAA 1449
 Db 1620 REDQ-----EQLOEETKRLNRQLOARSSIDNENIVSE--REBVLLEELALKQOLSLAG 1670
 QY 1450 EERL--EIKRDYFELYQTANTNLVBEKLETPLOADHEEDSIDRSEEMETIVLGEKLER 1507
 Db 1671 REKICELRNSS---TQONGNENGQVE--EQTFKEKEDRKPEDVPELISME--- 1720
 QY 1508 NOYLLERLOEBKLELSKLEITLOKEMETSVLKDD-----LOOKLESLESENTILKE 1559
 Db 1721 -RYALQAN-----NRLKILLEVYKTAAYABETIGRHVGLIDLRSSSOSASILWRS 1773
 QY 1560 NIDTTLK--HNSDOAQOLQK-----TQOELQKKNL---AIAAS 1593
 Db 1774 EAAVASKVSCHEHTVNTVDESIPYSGSDMPRNDIMMSKYTBEGELSORLYRSGFAGT 1833
 QY 1594 DNCPTIOE-----KETSADCVHPEELKILLTTEELHOKTNEOKTL 1633
 Db 1834 EIDPNEEELMNTISSRLQAAVEKTLBAISETSLHAYVQTELMRESFRQKQATESTL 1893
 QY 1634 ---LHEKNEL--EQAOVELKCEVEHLMKSMLESSESSLESLOHEKHD---TEOOL 1679
 Db 1894 KCOEELRERLHESRARBOIAVELS--KAGGVIDGYADETLFEROIOETKDIIIDRLDEL 1952
 QY 1680 LALKQOQVYTOEKKELQOTHEHILTAEVNHLKENILELGLNFKNQAOQTKKQOCLINERK 1739
 Db 1953 LCASNRLQLELEAQOQIOBERELLSHQKRAMK--AAG-----PVQOOLLOETE 1999
 QY 1740 ELBQSOHRLOCEIEELMKSILKDRSALFTLKSEBQVIVNLQNMENVMLEMBELNSORT 1799
 Db 2000 KLMKEKLEVOCAEKVRDLOKQVKALE-----IDVEQVSR--FTELQEKNTIELM 2049
 QY 1800 VIAERDQODDLRESEVEMSIETODDLRKAQOALQOQKQVOELTQISOISVL----- 1849
 Db 2050 DLKQOQALEKQLEKRRKRLDQALDREHERVFOQ--ETQLOEQOLKAVPRFOPISEHO 2107
 QY 1850 QEKISILENQMLYNATVRETLSEBDLNOQSKQHLSEFSETILTSIKKEKAPAL----- 1902
 Db 2108 TREVELOANHLREKTDKSELLSKREOLORDIOENNEETEKLEFVRLEQOLALVEDRKH 2167
 QY 1903 ---EQAKRDADARKITIDITEK--ISNIEBOLLOQATNLKETEYERSL--IQCKE 1952
 Db 2168 FGAVEAKPELSLEVOLOARDALDKREKETTNEBOLQOFREELKNKEEVOOLMOLEI 2227
 QY 1953 QALANTHELRETLKSKDLAGKME-----QERDEANKYALATEKMSLEB 1998
 Db 2228 QKKESTIRLOLEQENKLFKQDMKELGLAIKESDAMSQODQVHLEGRKQOITIQEKEVID 2287
 QY 1999 QINENVTTLKBEQGEKTFYIQRPSKQOSSOMEELRESLTKDLOLE-----E 2047
 Db 2288 QINBOVTKIQ-----QOLKITTDNKVIEEKNELIDLELOETQICLMSDQECVARN 2346
 QY 2048 AKKEISEATNEIKNLAKISSLEEBEITLQNASILNVAVERENLRHSKQOLYSE---LEQ- 2103

```

Db 2337 REEBEIOINVEIEKLOELANIGCOTSMANHSLSB---EADSLKQOLVIAEKLALQO 2393
QY 2104 -----LSLTLKSDHAFQ-----SKREDAVNRKIASLAE-----IKILTK 2141
Db 2394 VETANEEMTFMKVNFKNLQTLQELFSLKRER-ESVERKQIPENSVAVALDHLJSK 2452
QY 2142 EMDER-----RDSKESLOBSHLSSE-----LCTYTELOMLKQOEKEDINNKIAEKVK 2190
Db 2453 DKPELEVVLTDALKSLSLQTYFNSFEENGKSIINLETLLQLESTVSANDLELQCYK 2512
QY 2191 EV-----DELLQ-----HLSLKEQ-----LDQIOMELRRE----- 2216
Db 2513 QIKMOEGOFETEMLOKKIYNLOKIVEKVAALVSOILEFAVEYAKFCODNOTISSE 2572
QY 2217 -----KLRYELCEKMDIMEKETSIVRLMONEP-----QOEDBAVERMDILES 2260
Db 2573 PERTNIONMLQREDELSDISALTLRISELESQVYEMHTSLILEKEQVEIAEK-NVLEK 2631
QY 2261 RNO---EIQELMEKISAVYSEQHTLSSLSSELOKTEHAKHCMINIKSLSSTLSRSGS 2318
Db 2632 EKKLELOKLEEGNEKKQREKRSPODVEVLKTTLELFH-----SNEBSGFENLEA 2685
QY 2319 LQTEHVKLNTQLOTLNKFVYVRYTAAYKEDH--SLIKOYE--KDLAEOK-----RH 2367
Db 2686 LRASVATKALASVKEKAEKLOEBLYKETINMTSLQKDSQVRDHLAEAKESLILEKE 2745
QY 2368 DELRLQ---LOCL-EQHGKRWSDASAE---LKFCELEFINELLEFRANITQSVODDFS 2419
Db 2746 DETEVOESKKACMEPELPIKLSKSIASQTDGTLKISSNQTPQILVKNAGIQINOSBGS 2805
QY 2420 --EVOVFLNQVSTLQEELEHKKGMQWLEFPGDLHVDAKLSBMOENRIAST----- 2473
Db 2806 SEETVELTISO-----FTEKIEKMOELHA-----AETLDMSHISSETETLK 2846
QY 2474 -----IOLTKR---LKAVVO-----SKI-----OREIT-----VY 2496
Db 2847 REHVAVOQLKEEGTLLKAVIQCRLSKREGSSIPELASHDAVQTRICSDSGSDMGQGIY 2906
QY 2497 LNO---PEAKLOEKKEON-----KELMRMEHGPSA-----SVM 2528
Db 2907 LTHSGFDIASEGGESESATDSPPKKIKGLLRVHNEGQVLSLTPESPYSDEGDSIQ 2966
QY 2529 E-----EENARLLGILKTVODESKLOSRIKMLENELNLWKDAMH-----KGEKVAI 2576
Db 2967 QVSEPMLEERKAYINTISSLDLTKMO-----LQREAEVYDSSOSHSFSDMRGELLLA 3021
QY 2577 LQDKLL-----SRNAEFLNMQVYLTKRKQDNLQAMK----- 2609
Db 3022 LQVFLERBSVILAAFTLALGTDAVGLNLEQRIQOGVYEQAMCLOKADRRS 3081
QY 2610 ---EIEMLQKNAKAVPYKEID-----NKTK-----VAKIEBKIKYSAT 2650
Db 3082 LLSEIOALHMOANGRKITLKREOSEKPSQLELEYNIOKOSOMLEMOVELSMMK-DRAT 3140
QY 2651 DOEIAVYKSCLEDEKEGRLKEELRRAQADNDTTCVPKYOXKASTPVCGGSGSIGVQ 2710
Db 3141 E-----LOBOUSSKEMVAALKSELAQTKLELTLLKQKHKL----- 3179
QY 2711 STAMVLQSEKALARELSHYKKRYHHLSTRMSSSEDR-----KTKRASKDAHSHT 2762
Db 3180 -----ELEAFRLVYKDKTDVHLLNDTLASBQKKSRLQWALEKRAK-----L 3223
QY 2763 GSHRSGPHKTEYTRHG-PVTPPERSEMPSLHSGPKKSESSTKRVANSNBEITYSOLVMS 2821
Db 3224 GRSEERDEKELEDLKFSLESQKORNLQNLLEQOKOLNSQOKIESQRLMYDAQLSSE 3283
QY 2822 PGKGMKHHLISPSKVGHLKKRRLSPNREMPQTHAVISPGTGLHKNLTESTLFDONLSSP 2861
Db 3284 QGRNLEQVLLSESEKVRKRREMSSTLDRERELHAOLQSDGIGQSRPLPSDELLKELQO 3343
QY 2882 C--KOQKVOENLN--SPKGLFEDVSK 2904

```

```

Db 3344 LEEKSRIVELLNTEPEKYNKIDSLQTR 3369

RESULT 3
US-10-171-311-2
Sequence 2, Application US/10171311
Publication No. US20030087270A1
GENERAL INFORMATION:
APPLICANT: Schlegel, Robert
APPLICANT: Chen, Yan
APPLICANT: Zhao, Xumel
APPLICANT: Monahan, John
APPLICANT: Kamatkar, Shubhangti
APPLICANT: Glatz, Karen
APPLICANT: Ganuavarapu, Manjula
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
FILE REFERENCE: MRI-035
CURRENT APPLICATION NUMBER: US/10/171,311
CURRENT FILING DATE: 2002-06-12
PRIOR APPLICATION NUMBER: US 60/298,159
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/298,155
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/335,936
PRIOR FILING DATE: 2001-11-14
NUMBER OF SEQ ID NOS: 238
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 3907
TYPE: PRT
ORGANISM: Homo sapiens
US-10-171-311-2

Query Match 6.9%; Score 1013; DB 9; Length 3907;
Best Local Similarity 19.1%; Pred. No. 7.4e-31;
Matches 691; Conservative 637; Mismatches 115; Indels 1182; Gaps 148;

QY 181 KKEGNRRHYGETKNDHSSRHTTIFRMIVESRDRNDPTNSECDAVNVSHLNVDLGAS 240
Db 34 KKQKKR---KTSSEKHDSVAHHDNLNIDQOCNEKYNINSQREVESTVPESTIMRTLSHG 90
QY 241 ERASGTAGVRLK-----EGCN--INRSLEFIIQVYIKLSQDQAGFIYRPSKTLR 291
Db 91 EITSHQGSFVSELESEISTADDCSSEVNGCSFV-----RTGKPTN 132
QY 292 ILONSLGNATVICTITPVSFDETLSTLOFASTAKHVNRTPHYNEVLDEALALKRYRK 351
Db 133 LIREBERG-----VDDYS-----EGQADOSPHTLEMMESLAKQH-- 169
QY 352 EILDKQOLENLESSESTKA-----QAMAKEH--TOLLAETKQHLKEREIRIMILT 401
Db 170 EIEELNRLLEMRVYTGREGLOQLOEFPAJKORGIITQTLANQARREDELMREFL 229
QY 402 NIYVASOESQODQVRKRRRYTAPKQIONSLSHAGVSD-----FDMLSR 447
Db 230 EL-----TEOSQLOIQFOLOASFTLRNTHSSTADLLOAKOQILTHQOOLEEODH 282
QY 448 LPGNFSK-----KAFSDMPSFEIDSDVCTEFSDFPDALSMWDSNGIDAE 493
Db 283 LLEIYQKKEDFTMOISLQEKIKIYVEEQDKKYNNSKKELOEKETIIEELNRTIIEE 342
QY 494 ---WNLSKVTHEREKT--SLHQSMIDFGQ-----HSDSVQPHDSSKENOLOYLPRDS 540
Db 343 KKTLELKDKLTTADKLGLLEOEOIYQKQOEIKNMKLELTNKKQERQSEETIKOLM---- 398
QY 541 GDMAECR-----ASFKEIITS-LQOQLOSK-EEEKKRLVOSFELKIALDEOL--SVKA 591
Db 399 GTVELOKRNNHKDSQEFEDIVQRMEOETQRLQELRAELDEMYGQOIVOMQOELIRHMA 458
QY 592 KMLEAVTNSR---EHSINAEVQTDVEKEVVRKEMSVLSDSGYNASNSDLQSSVSGRRLS 648

```



```

Db 459 QMEKKTFRHKEMENALNSYSNITVNEQIK-----LMNVAINEINIKIQTNSQOKETLK 513
QY 649 SSHDCIEHRKMLBOKIYDDEEFIEINENKKSSENDKOKSSPDQF-----ESIQCEALMA 703
Db 514 EELGILIEKAKALDOLDEEELSPFSREQIARAKOTABESKINEAKHKLSTVEDIKA 573
QY 704 EKANALE---ELAL-----MRDNFIILIENTLREIADLER-----738
Db 574 EIVASESRKLELKHAEVYNYKIKLEMEKENAVLDRAEOP- AELERLKTOLLS 632
QY 739 -----SLKE-----NOETNEFEL- EKETQREHE 761
Db 633 HEEELSKLEDELEIHRINIEKKNIGIHYKQOIGDIGNEMSOIKETMOEKNLITKQ 692
QY 762 AOLHIEGSLKLEVENAMYNQNL-----DEETKRIKLE-----QEIOL 803
Db 693 NQILLETISKLOLOS- LNSKSEEMTLOINLOKEIETIROEKEKGTLEQEOVELQ 750
QY 804 -----AELRRADNLQKVRNFDLSVS-MGDSKLECEIFOLKQSLDAEAV-----849
Db 751 KTELLEKQMKENDLOEKFAOLFAENSILKDEKTLIEDMLKHTPYQOEERLITFDSIK 810
QY 850 ---TQDA---QRECSFLRSEN-----LEKEMDTSNWTNOKKASALEKOLETEKSN 898
Db 811 SKSDSWMEKEIETILEENEDLKQOCIOLEMEIKORNTSFARKEFVNYQELQOEYAC 870
QY 899 YKKMEADLO-----KELO-----SAFNEINYL-----921
Db 871 LKATKDLDESKNQEILEYSKILKALNEHLORINPTVYKMSVDEDEKTPAETLEM 930
QY 922 GLLAGKVPRLD-----SRVLEKKEVSFOLEKALEKNALENEVTCISEYK---970
Db 931 GEVEKOTTEIEMERLETKREKTELISORTLSIDSLQKQKEISFLNEEYKSLQKEQOV 990
QY 971 -FLUNEVECKLN-----QISKASEIMLKQEB- EHSASTISK-----1006
Db 991 SLKRELEIITINNRANVOSCTOVSSLLDGVVMTSRAGESSVKNKSPGESSKIMV 1050
QY 1007 -----QEIIMQOSEO---ILQITDEVTHQS---KVQOTEQOYLE-----MKKMHDLF 1050
Db 1051 EDKVSFNMTVGEESKQEOILDLHPVTKESSLRATQPBENDKQKELAVLKSQNDL- 1109
QY 1051 EKYIRNKSAB-----DLRE- MENIKGT-----MESVEYKIADTKH 1086
Db 1110 ---RLOMEQORICLSLVYSTHVQVREYEMENEKDALCSIKELIFPAQEKIKELQIKH 1165
QY 1087 ELE-ETIR-----DKEOULH-----EKKYFPOAMOTIFPTPLSDSLPSPKV 1128
Db 1166 QLEIQTKTQETGDEGKPLHLILGKLOKAVSECSYFLQTLCSV-----LGEYTPALKC 1220
QY 1129 EGSNODPIELNDY-----HNLTALATERNNIWCLET-----1160
Db 1221 EVNAEDKENSQDYISENEDELQDYRYQODQENMHTLLNKVTEENKVLVYQTRLSKI 1280
QY 1161 -----ERNSLKQOV-----IDLMTQLOSOLOQSIKESDLQKPR 1193
Db 1281 WGOQTQMKLEFEGEBENPKETEFELSHSQMTNIEDVNHKSKLSLDLEKTKLEQOV 1340
QY 1194 QDLE-----EGEVNLLLEMLLKGLH-----TDSOLST 1221
Db 1341 QELIESLSSLOQOLKEEQVYEAHICLOKRLQAVSESYPPSLPDSVYITTESDQKRTA 1400
QY 1222 -----EKLQLENLEVERKLOTLQOEMKNITI-----ERNE 1251
Db 1401 YPGSCVKKNIDGTIEFSGEVKKEETINIVKLEKQYOEOQEEVAVYVMSIAFAQOHE 1460
QY 1252 L-----QTNFEDLKAHDSLKODLSSENIQOSIETODELRAAQ 1288
Db 1461 LSHISGKENTASSKQAHAAVQOEOHYFNMKLSODOIGQPTREYD-- VKFEERKPPIS 1518
QY 1289 EELREO-KOLVDSFRQ-----QLDSCSVGIS---SPNHDVAVANO---EKVSGEYVNS- 1333

```

```

Db 1519 KELGEHGEKILLSNDPDHIDPESKCVLTIISEMFSKOTFIYROSITHDEISVSMDSAR 1578
QY 1334 ---LQSMNGERDELOTSOKALVSELELR-AHKSVEBENLEITKNGLEKELLGKS 1389
Db 1579 QLMNEBOLBEMROELROYOEHOATELLRQAMRQME-----RQ 1619
QY 1390 ESEVELKSMLENKEDNNKLEQAEYSKKNQFSLSEVSGSOKLYDEIYVLAQULKA 1449
Db 1620 REDQ-----EQLOEIKRLNRQLOAQSSIDNENYSE---REVRILLEELALKQSLAG 1670
QY 1450 EERL--BIKRDYFELVQANTNLVEGKLETPLOADHEEDSIDRSEMEIKVGLER 1507
Db 1671 REKICELLRNSS---TOTONGNENOGEVE---EQFERKELDRKEDVPEITISNE---1720
QY 1508 NOYLLERLOEKEKLSKLETLQKEMETSVLLDD-----LOQKESLEBENITIK 1559
Db 1721 -RVALQAN-----NMLKTLILEVNTVTAABETIGRHVLGILDNSKQSSASLWRS 1773
QY 1560 NIDTTLK---HNSDTQAOLOK-----TOOELOLAKNL---AIAAS 1593
Db 1774 EAEASVSKVCHHEHTRTYDTSIPSYSGDMPRNDINMKSVTBEGNELSGRLVRSQFAGT 1833
QY 1594 DNGPTIOE-----KTSADCVHPLEKILLLTHELHQGTNEQEKI 1633
Db 1834 EIDPENELMLNITSRLQAAVEKLELBAISTSSQLEBAKVTQTELMRESFRQOZATEST 1893
QY 1634 -----LHEKNEL- EQAQVELKCEVEHLMKSIESKSSLSLOHEKHD---TEOOL 1679
Db 1894 KQOELREHRESRAREQOLAVELS-KABEVITGYADEKTLIFRQIOEKTDIIDRIEDEL 1952
QY 1680 LALKQOVVTOEKELOQTHEHTAEVDHLKENEILGLFNKNAQOKTKTEQOCLINENK 1739
Db 1953 LCASNRLQOELAEQOQOQOEBRELLSRQEKAM--AEAG-----PVEQOLLQETE 1999
QY 1740 ELQOSQRLQOCELEMLKSLKQKESALETKESQVYININQMEMVMLEMELKNSQRT 1799
Db 2000 KMKKEKLEVOQQAQKVRDDLOKQVKALE-----IDVEQVSR-FTELQEKNTIELM 2049
QY 1800 VIAERQLODLDRESEVMSIETQDDLKQAEALQOQKDKVOELTQOIS-----VLQ 1851
Db 2050 -----DLRQONALEKOLEKMRKRLDLQALDREHERVFRQ 2085
QY 1852 KISLENOMLYNAVYKTELSEKDLNOSKOHLESEITLSLSIKKE-----FALBO 1904
Db 2086 EIOKLEQOUL---KVYPRFOPIS-----HQTREVEOLANHLKRETKCSLELTSKQ 2134
QY 1905 AEKQKADAARKTIDIEKISNIEEOULLQOATNKETLYERE-SLICQKQOLAMT--EHL 1961
Db 2135 LQDIOERNEDEIKLEFRVARELQALLVSADTFQVEDRKHFGAVAKPELSLEVQLOAE 2194
QY 1962 RETLKSKDLALQKMEQO---RDEAANKYIATLEKSSLEEOINENVTTLKGEGERKEF 2017
Db 2195 ROAIDRKEKEITNLQELQFRELLENKNBEVQQLMQOELQKKESTTLQOELQONKLF 2254
QY 2018 YLQRPKQOOSQMEELRSLTKDLQLEBAF-----KEISEATNEIKMLJAKISL 2069
Db 2255 -----KQDMERLGLAIKESDAMSTQDOHVLEGRKFAQITIOEREVEIIDOLNEAVTL 2304
QY 2070 EEEILONASILNEVASERENLRHSKQOVLSELQSLTSLKSDHAFQSKRREDEAVNKT 2129
Db 2305 QOOL-KITTDKVIIE-----KNELTRDLETQIECLMSQOECYKRRREBEITROLN 2355
QY 2130 A-----SLAEIKITLTKEMDEFROKSESLOQOSSHISELCTYKT- 2169
Db 2356 IKLOQELANIGOKTSMNASHLSSEADSLKHQDLVIAKLLALQOQVETANEMFMKNV 2415
QY 2170 -----ELQMLKQOKEDINNKLAERYKEVD-----LIQHSLSKELOLDQIOM 2211
Db 2416 LKETNFKMNLQLOELFSLKRESEV---EKIQSIPENSVAVAIDHLSKDKPELEVLVT 2470
QY 2212 E-----LRNEK-LRNECEKMDIMEKETSVAL-----MONQ 2245
Db 2471 EDALASLENQTYFKSPSENGKSTINLETRLQLOESTVSAKDLLELTQCYKQIKDMQOEGQ 2530

```

```

QY 2246 OEDDVABMDLIESRNOEIOEIM-EKISAVYSEQHTLLSLSELOKET--EAKHCHL 2302
Db 2531 FE-----TEMLQKRYNLAKIVEEKVA-----ALVSIQLEAVOEYAKFCOD 2573
QY 2303 N-----IKE-----SLSTISRSFGSLQTEHVKNLQTLQTLNKKVY 2341
Db 2574 NOTISEPERTNIONLQREDELSDISALTLRRISELESQVEMHTSL--ILLEKOEVI 2631
QY 2342 RTAVKEHSHLKDYEKLAEBQKRDHLRLOQLCEQHGRKWSASASELK----- 2393
Db 2632 AEKNVLEKEKKLLEQLKLEGNKKQREKE-----KKRSPQDVEVLTTELPHS 2681
QY 2394 -----FCEIEFL-----NELFKRANLIQSDODESEVOYL 2425
Db 2682 NEBSGFPELALRAESVATRAELASYEKAKELQEBLLVETN--MTSLQDLDSQVRHL 2740
QY 2426 NOVSTIQ-----EELHNKG----- 2441
Db 2741 AEAKELSLLEKDETEVOESKKAQMFPLPIKLSKIASQDTGLKISSNQTPQILVK 2800
QY 2442 -----FMQMLEEGDLHVDAKKISEGMOENRIAST----- 2473
Db 2801 NAGIQINLOSECSEEVTEIISQFTEKIEKMOELH-----AEILDMESRHISLETELKR 2855
QY 2474 -----IOLLTKR--LKAIVVO-----SKI-----OREIT-----VYL 2497
Db 2856 EHYAVVOLKEECGTLKAVIOCLRKESGSIPELAHSDAYOTREICSDSDSGMDGQGYL 2915
QY 2498 NO--FEAKLOEKKQN-----KELMRMEHHGPSA-----SYME 2529
Db 2916 THSOGFIASGREGESATDPPFKIKGLLRAVHNEGMOVLSTESPYSDGDEHSLQ 2975
QY 2530 -----EENALLGLIKVVOESKTLQSRKIMLEMLVYDDAMH-----KGGKVAL 2577
Db 2976 VSEPMLEKRAYINTISLKDILITMO-----LOREAEVYDSSGSHESFSPWRGELLAL 3030
QY 2578 QDKLL-----SHNAEAEILNAMOVLTKRKODNLOAMK----- 2609
Db 3031 QOVFLERSVLLAARFELTALGTDAVGLNCLDEQRIOEGVEQOAMECLOKADRSL 3090
QY 2610 --EIBNLOKNAVAKAVPYKEID-----NLTK-----VKTMEKIKYSKATD 2651
Db 3091 LSEIOALHMOONNGRKITLKRQDESEKPEQOELLEYNIOQKOSOMLEMOVALESMK--DRATE 3149
QY 2652 QEIVAYKSCLEKDEKGLRLKEELBRAQADNDTYCVPKDKQKASTFVTGCGSGCIYOS 2711
Db 3150 -----LOQOLSEKKNVAELKSELAQTLLETLTLKAQHKLK----- 3187
QY 2712 TAMLVLOSEKALERELESHYKKKYHHLSTRSSSEDR-----KTKAKSDAHSHTG 2763
Db 3188 -----ELFAFLRELVADKDEVHLLNDTLASEOKKSRLEOMALEKKAK-----LG 3232
QY 2764 SSHHSGPHKTYETHRG--PVTPERSEMPSLHLSGPKKSSSTKRVVSPNRSSEIYQVLPSP 2822
Db 3233 RSEBRDKELDLKPFSLSSQONMQLNLLLEQKQLNNEQOKLESQRMLYDQALSEBO 3292
QY 2823 GKTGMHKLHLSFSPKVGGLHKKRALSPNSEMPTQVHVISGKTGLHKNLTESTLFDNLSSPC 2882
Db 3293 GRNLEQVLESEKVRIRIEMSTJDRERELHAQLOSSDGTGOSRPLPSEDLKELQROL 3352
QY 2883 --KOQVOENLN--SPKGLFDVYKSK 2904
Db 3353 EKHSHRYEELNETEKYKLDLSQTR 3377

```

RESULT 4
US-10-171-311-6

; Sequence 6, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan

```

; APPLICANT: Zhao, Xumel
; APPLICANT: Monahan, John
; APPLICANT: Kamathar, Shubhangi
; APPLICANT: Glat, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoersli, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 3925
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-171-311-6

Query Match 6.9%; Score 1013; DB 9; Length 3925;
Best Local Similarity 19.1%; Pred. No. 7.4e-31;
Matches 691; Conservative 637; Mismatches 1115; Indels 1182; Gaps 148;

QY 181 KKEGNRHGTEKMDHSHRSHTIFRMIVESDRNDPTFSENGDGVVSHLNLVLAGS 240
Db 34 KKQKKR--KSSSKHVSAAHDLNDIDOSCENEMYINSQVESTVPIESTIMTLHSG 90
QY 241 EPASGTAGEVRLK-----EGCN--INRSLFIIQVYIKLSDQAGGFYWRDSKTLR 291
Db 91 EITSHGOGFSVLESEISTTADDCSEVNGCSFVM-----RFGKPTN 132
QY 292 ILQNSLGNKAVIITCTPVSFDETLSTLOFASNAKHVRYNTPHVNEVLDEALLKRYRK 351
Db 133 LREBERG-----VDSYS-----EQGADSPHLEMMESLAGKH-- 169
QY 352 ETLIDKLOLENSSSEKRA-----QAMAKEH--TOLLAETIKOLKEREDEIRIWL 401
Db 170 EIEELNRELEEMRYVTGTEGLOQLQOFEALYQORGIITQULANLOQARREDEIRIWL 229
QY 402 NIVVASSQSDQVKKRRRYTAPGKIQNSLAHGSVD-----FDMLSR 447
Db 230 EL-----TEGSKLQIQFOQLQASEFLRNSTHSSSTAADLLQAKQILTHQQLLEQDH 282
QY 448 LPGNESK-----KAKFSMPSPFEIDDSCTEFSDDDALSMDSNGIDAE 493
Db 283 LLEDYQKKKEFTMOISFLQEKIKYVEMEQDKVENSNNKEEIOEETIIEELNTIIEE 342
QY 494 --NVLASKVTHREKT--SLHOSMIDFG-----ISDSVQFIDSSKENLOLVLPRDS 540
Db 343 KKTLELKOKLITADKLIGELQOIVQKNOEIKNMKLELTNSQKROSESEELKQML-- 398
QY 541 GDMACR--ASFEKEITS--LQOOLQSK--EEEKKELVOSFELKIALEBOL--SYVA 591
Db 399 GTVEELQKRNNHDSQFETDIYQRMQEQTRKQRAELDEYGGQIQVOMKELLRIQMA 458
QY 592 KMLEVNTSR--EHSINAQVOTDYEKVEYKREKSVLQDGSYNASNDLOOSSVDGKRLS 648
Db 459 QMEEMKTRHKGEMENALRSYNTIVNEOIK-----LNNVALNELINIKIOPDINSOKELIK 513
QY 649 SSHDECIEHRKMLEQKIYDLEEFIEENLNKSSENDOKSSEDFM-----ESTQLOEALMA 703
Db 514 BELGLILEKCALQORQDLVLEELSFSSBOQORAQOTTAEDQSKINEAKHSSTVEDIKA 573
QY 704 EKANALE--EIAL-----MRDNFDNIILENETIKREIADLER----- 738
Db 574 EIVSASESRKLELKHAEVNTYKIKLEMLEKREKNAVADRMAESOE-AELERLRTQOLFS 632

```

QY 739 -----SKR-----NOETNEPIL--EKETOKENE 761
 Db 633 HEBELSKIKDELIEHRINIEKLNIGIHYKQOIGLOENEMSOIKIETMOFEKDNITTKO 692
 QY 762 AOIIEIGSKIKVENAMYNOLIE-----DETETKILKE-----OELQ 803
 Db 693 NOILIEISKIKDILOS--LVNSKEEMTLOINELOKIEIILROBEKEKGLIEBOVLOL 750
 QY 804 -----AELKRADNLOKVRNFDLSVS--MGSEKICEEITFOUKOSIDAEAV----- 849
 Db 751 KTELLEKOMEKENDELQEKFAOLEAENSILKDEKKTLEDMLKITHPVSOBERLIFDSIK 810
 QY 850 --FRDA--OKECSFLSEN-----LEKEKEDNSNMVNOKEKAASLEKOLETEKSN 898
 Db 811 SKRSDWEKEIETLIEENEDLKOQOICLNEEIEKORNTISPAKEFVNYOELQEIYAC 870
 QY 899 YKREMDLO-----KELO-----SAFNEINYLN----- 921
 Db 871 LKAVKDLEDSKNOKELEYKSKIKALNEELHQRINPTVAKMSVDEDEKTFVAETLEM 930
 QY 922 GLLAGVPRLL-----SRVELEKVSFESKOLEKALEKNALENYTCLSEYK-- 970
 Db 931 GEVEAKDTTELMEKLEVTYKREKLELSQRLSDLSBOLKQKHGEISFLNEEYKSLQKEKQY 990
 QY 971 -FLPNEVECLKN-----QISKASEEIMLKQEG--EHSASITSK----- 1006
 Db 991 SLKRELEIILINNRAENVOGSDTOVSSLLDGVTMTSRAGEGVSXVNSPGESEKIMV 1050
 QY 1007 -----OETIMOROSEO--TLOUTDEVTHTOS--KVQOTEOBYLE----MKRMDLFL 1050
 Db 1051 EDKVSFENMTVGESEKQEOILDLPLSPVTKESSLRATQPBENDKLOKELNVLKSEOND 1109
 QY 1051 EKYIRNKSSEAE-----DLRE--MENLKG--MESVEKIKADTKH 1086
 Db 1110 -----RLOMEKQORICLSIVSTHYDQVREYMEKOKALCSLKEELLPAQEKIKELQKH 1165
 QY 1087 ELE-ETIR-----DKEDOLH-----EKYFFQAMQITFTPTPLSDSLPSKLV 1128
 Db 1166 QLELOTKMQOTGDEGRPHLLIGLKQKAVSEBSYFLQTLGSV--LGEYTPALKC 1220
 QY 1129 EGSQOPFIEINDY-----HNLALATERNNIMVCEY----- 1160
 Db 1221 EVNAEDKENSQDIYISENEDELQDYREYODFOENMHTLLNKVTEENKLLVLQTLRSKI 1280
 QY 1161 -----ERNSLKEOV-----IDLNTLOSLQOASIEKSDLOKPK 1193
 Db 1281 WGOOTGMKLEFEBENPKRETEFELSHSQMTNEDIDVNHKSLSSLOLLETKLEQY 1340
 QY 1194 QODE-----BGEVALLMEELLKGLH-----TDSQLOST 1221
 Db 1341 QELESLSLQOOLKETEQYEAIEHLOKRLQAVSESTVPPSLPVDVVTESDARQM 1400
 QY 1222 -----EKLOLENLEYTEKLOTLOEEMKNITI-----ERNE 1251
 Db 1401 YPSSCVKKNIDGITEPSGEGVAKETIVYKLLKQYOEOLEEBEYAKIVMSMAFQOTE 1460
 QY 1252 L-----QTNFEDLKAHESLQKODLSENIQOSIETODELRAAQ 1388
 Db 1461 LSRISGKENTVASSKQAHAVQOQOEHYFNEMKLSQDQIGQTFEYVD--VKFKEEPFLS 1518
 QY 1289 EELREO-KQLVDSFRQ-----QLDCSVGIS--SPNHAVANO--EVSISGEVNS-- 1333
 Db 1519 KELGEHKEKILLNSPDHPIRESDCVLTISEMFSSKDKFTFIVROSITHDEISVSSMDASR 1578
 QY 1334 ---LOSEMLNGERDELQOTSCALVSELELR--AHVKSVEBENLEITKKLNGLEKELIGKS 1389
 Db 1579 QLMNLEBQLEDMKQOELVQYOEHQOATELLROAHMRQF-----RQ 1619
 QY 1390 ESESVLKSMDENLKEDNNKLEQAEEYSKBNQFSLIEEVSQSGSKIVDEIVLKAOLKAA 1449
 Db 1620 REDQ-----EQLOEETIKRLNRQLOARSSIDENLVSE--REVVILLEELFALKQJSLAG 1670

QY 1450 EERL--EIKDROYELVQANTNLVEGKLETPLOADHEEDSIDRSEEMEIKVIGELER 1507
 Db 1671 REKLCEELRNS--TQTONGENOGEV--EQFFKEKEDRKEDEVPPELISNB-- 1720
 QY 1508 NOYLLERLOEKELELSNKEILLQKEMETSVLLKD-----LOOKLESLENIILKE 1559
 Db 1721 -RYALOKAN-----NRLKILLEVYKTAAYBETIGRHVGLIDLRSSQSQSSASLIMS 1773
 QY 1560 NIDTTLK--HNSDQOAOLOK-----QOELQOLAKNL--AIAAS 1593
 Db 1774 EAAVASKCVHEBHTVDESIPSYSGSDMPRNDIMMSKVTEEGTELQORLRSQFAGI 1833
 QY 1594 DMCPIQOE-----KETSADCVHPEEKILLTEELHOXTNQOEXL 1633
 Db 1834 EIDPENBELMNTISRLQAAYEKLLEALISSTSQLHAYQTOELRESPROQAEATESL 1893
 QY 1634 -----LHEKNEL--EQAVELKCEVEHLKMSMESSSLSLELQEKHD--TEQOL 1679
 Db 1894 KQOEBLERERLHESRARQOAVELS--KABGVIDGYADEKTLFEROIQETDIIDRLQOEL 1952
 QY 1680 IALQOQOVYTOEKLELOQTHEHLTAEDHLKENEILEGLNFKNEAQOKTKEQCLINENK 1739
 Db 1953 LCASNRLOELBAQOQOQOIERELLSRQKEMK--ABRG-----PVQOLLOEIE 1999
 QY 1740 ELBOSQRLQCEIEELMSLKDESALETLKESEQKVINLQEMENVMLEMEELKNSQRT 1799
 Db 2000 KLMKEKLEVOCCAKEYRDDLOKQVKALE-----IDEEQVSR--FIELEQEKTEML 2049
 QY 1800 VIAEROLODDLRESVEMSLTEODDLRKAQOALQOQKDVQOELTSQIS-----VLOE 1851
 Db 2050 -----DLRQOQALQEKQERKFLDQALDREHERDVFQO 2085
 QY 1852 KISLENOMLYNAVYKETLSEKRDNLNOSQOHLFSEITLSLSKEKE-----PALQ 1904
 Db 2086 EIOKLOQOL--KVPRFPQISE-----HQTREVEQLANHLKKTOKCSLELSKQ 2134
 QY 1905 AERKQADANKITDITEKSNIEBOLLOQATNLKETLYERE--SLIOCKEQLALNT--EHL 1961
 Db 2135 LQRDIOERNEIEKLEFRVARELOALLVSDTPQKEDRKHGAVAKKELSLQVLOQE 2194
 QY 1962 RETLKSQDALQKMOE--RDEANKVYALPEKSSSLEQOINENVTYKKEGEGKEF 2017
 Db 2195 RQALDKREKITLLEQLEQFRELKNEVQOLMQLEIQKKESTRLOLEQENKLF 2254
 QY 2018 YLQRPKSQOSSQMEELRESLTKYDLOLEAE-----KEISEANIEKNTLAKISSL 2069
 Db 2255 -----KDMERKIGLAIKESDAMSTQDQVHLFGKFAQIOEKREVELOLQNEQVTXK 2304
 QY 2070 EEBLLQNASILNEAVSERENLRHSKOOLVSELBOLSLTLKSRDHAFARQSKREDEAVNKI 2129
 Db 2305 QOOL--KITTDNVIEE-----KNELLIDLETQOELCMSDQECYKRNEEBEIOLENY 2355
 QY 2130 A-----SLAEIKITLKENDEFDSKESLOEBSHLSBELCTYKT-- 2169
 Db 2356 IEKLQELANIGQKTMNASHSEADSLKHQDVIYAKLAEQOEVARENEBETMKNV 2415
 QY 2170 -----ELOMLQOQKEDINNKLAEKYKVEDE-----LLOHLSLKQDLOQIM 2211
 Db 2416 LKETNFKMNQLOELSLKRESV-----EKIOSIPENSVNVAIDHLSKDKLELEVLT 2470
 QY 2212 E-----LRNEK--LRNYELCKMDIMEKETSIVLT-----MONEPQ 2245
 Db 2471 EDALKSLQENQTYKFSBENGKSIINLETRLLQLESTVSAKDELQCYKQIYQIDMOBQO 2530
 QY 2246 QEBDDVAERMDILESNOEIOELM--EKISAVYSEQHTLLSSLSSELQKEN--BAKHCM 2302
 Db 2531 FE-----TEMLOKRTIVNOKTIVEKVA--ALVSQIOLEAVQYAFACQD 2573
 QY 2303 N-----IKF-----SLSTLSFSGSLQTEHYKLTQLOTLNKKRYV 2341
 Db 2574 NQRTISSEPERTNIONQLEDELDGSDISALTURLISELESQOYVEMHTSL--ILLEKVOEL 2631
 QY 2342 RTAAVVEDSHLINDYERKDLAABQKRHDELDLOQCLEQEHQKMSDSASEELK----- 2393

```

Db 2632 AKNVLEKELKLEQKLEBENKREKRE-----KKRSPDVEVAKTTTELEHS 2681
QY 2394 -----FCEIFL-----NELLEFRKANI IOSVODFSEVOVEL 2425
Db 2682 NEESGFNELEALRAESVATKALASYKAKAKLOEELIVAKETN-MTSLQDLISQVRDL 2740
QY 2426 NQVSTIQ-----FELEKKG-----2441
Db 2741 AEAKELISLEKEDETEVOESKKAEMFEPPIKLSKSIASOTDGLTISSNQPTILVK 2800
QY 2442 -----FMQWLEEGDLHVDAAKLSEGOENRIAST-----2473
Db 2801 NAGIOMLOESCESEVEYTEIISOTFEIKEMQELHA-----AELIDMESRIISTETELKR 2855
QY 2474 -----IOLLTKR--LKAIVO-----SKI-----OREIT-----VYL 2497
Db 2856 EHYVAVOLLEKECOTLKAIVOCLSRKESGSSIPELAHSDAYOTREICSSDSGDMGQGIYL 2915
QY 2498 NQ-----PEAKLOEKKEON-----KELMRMEHHPRA-----SYME 2529
Db 2916 THSGFVIASGEGSEESATDSFPKIKIGLLRAVHNEGMOVLSTESPYSDGDHSTIQ 2975
QY 2530 -----ENARLLGLIKTVODESKIKOSRIKMLENELIVKDMH-----KGEKVAIL 2577
Db 2976 VSEPMLEERAKYINTISLMDLITKMQ-----LQREAEVYDSQSHSFSFDMWRELLAL 3030
QY 2578 QDKLL-----SRNAEELNAMOVLTKKODMLQAMK-----2609
Db 3031 QOVFLERSVYLAARFETELTALGTDDAVGLNCLCLORIQIOEGVEYOAMELIQADRSL 3090
QY 2610 -----ETENLOKVAVAKGAVPYKKEID-----NLKTK-----VVIEMEKIKYKATD 2651
Db 3091 LSELQALHAQNGKKITLKREQSEKPSOELETVNIQKOSOMLEMOVELSSMK-DRATE 3149
QY 2652 QETAYLKSCDEKEEGRRLKEELRAQADNDITVCYPRKYOKASTEPVYTCGGSGIYOS 2711
Db 3150 -----LQEQLSSEKMYVAELKSELQOTKLEETTLKAQHKLK-----3187
QY 2712 TAMLIVLOSEKAALEREISHYKKKYNHLSRTWSSSEDR-----KTKAKSAHSSHTG 2763
Db 3188 -----ELEAFRLKVDKTDDEVHLLNDTLASQOKSRLOMALKEKAK-----LG 3232
QY 2764 SSHRSGPHTKETYRHG-PVPERSEMPSLHLGSPKKSSESTRKRVYNSRISYOLVMSR 2822
Db 3233 RSEBRKEELEDKFSLESOKORMLQNLLEQOKOLLNESQOKITESORMLYDQOLSEBO 3292
QY 2823 KGTGMHKLILSPSKVGLHKKRALSPNRSEMPYQVHISPGKTGLHKNLTESTLFDNLSSPC 2882
Db 3293 GRNLELOVLLESEKVRIRREMSSTLDRERELHAQLOSSDGTGOSRPLPSBDLLKELQOL 3352
QY 2883 --KQAKYQENLN-SPKGLFDVAKSK 2904
Db 3353 EEKHSRIVELLNETEKYKLDLQTR 3377

```

RESULT 5
US-10-082-830-260

```

; Sequence 260, Application US/10082830
; Publication No. US20030077604A1
; GENERAL INFORMATION:
; APPLICANT: Sun, Yonming
; APPLICANT: Recipon, Hervé
; APPLICANT: Salceda, Susana
; APPLICANT: Liu, Chenghua
; APPLICANT: Turner, Leah
; TITLE OF INVENTION: Genes and Proteins Relating to Breast Specific
; FILE REFERENCE: DEX-0249
; CURRENT APPLICATION NUMBER: US/10/082,830
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,802
; PRIOR FILING DATE: 2000-10-27

```

```

; NUMBER OF SEQ ID NOS: 282
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 260
; LENGTH: 2383
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-082-830-260

```

Query Match 6.7%; Score 993; DB 9; Length 2383;

Best Local Similarity 21.2%; Pred. No. 2.6e-30;

Matches 562; Conservative 472; Mismatches 929; Indels 690; Caps 101;

```

QY 370 KAOAMAKEEH--TQLLAEIK--QLHK-EREDRIHMLNIVASSQESQDQVRKRRRV 423
Db 227 KTOELKEAHERSOELLQOLKSGDLEKAELODRYTELSALTQSOQONEDKAKARE 286
QY 424 TWABGKIONSILHAGVSDFDMLSRPLGNFSKAKAFSDMSPFEIDDSVCTEFSFDALIS 483
Db 287 T-----VEILETNTHELMHEASL- 305
QY 484 MDSNGIDAEMNLASKYTHREKTSLSHSMIDFGOT-----SDSVQPHSSK 529
Db 306 -----SRNAQEKLSLOQYIKDITQVMVEEGDNIAQSGSHENSTEL-DS51 350
QY 530 ENOLOYLPKDSG-----DMAECKASFEKETSLOQOLQSKKEEK 569
Db 351 FSQFDYQDADKALLVNSVLTTRRQAOVDLQOLAGCOEA-----VNLLOQOHQWEEBEG 405
QY 570 KELVQSEFL-----KIAELEQOLSVKAKILEVNTNSRHSINAVOTDVEKEVVR 619
Db 406 KALNRQLOKLTGERDITLAGQVVDLQGEVDSLSKEBELLOKARE-----LR 451
QY 620 KEMSVLDGCVGNASND-----LQDSVDGKRLSSHD--ECIEHRKMLEQIVDLEEFIE 673
Db 452 QOLEVLEQEARLRRVNVLELOQDSQOGKEQOEELHLVRRERERLOEMLMGLA--- 508
QY 674 NLNKSSENDKQSSQODEMESTIOL-CEAIMAKANALEELALMDNPNITLLEMTLKE 732
Db 509 ---KQESLSLITLRELESSEHLEGEILROEQIEVTAALRAQASIAELSSSENTLKE 565
QY 733 IADERSIKENOETNEFEILEKETQKEHEAQLIHIGISLKLIVENAEMYNONLEDELETK 792
Db 566 VADLRAAVKLSALNEALDALDKVGLNQQLQLEENGVSYSRMAEAGARALVVD--- 621
QY 793 TKLKEQEIQLAEIRKRNADLQKRVNFDLSVSGDSEKLCCEETFOLOKOSIDAEVATRD 852
Db 622 -----LAEAEKR-----REALMEKNTHLQALQKAEAGAE 652
QY 853 AQKESCFRSENLLEKEMEDTSMYNOKEKAASLFEKOLETSESNVKKMKADLOKELOS 912
Db 653 LQADLRDIOEKEBETOKKLSER--HQQEAATTQLE-QLHQEAKROBEVLARAVOEKEA 708
QY 913 AFNEINYLNLGLACKVPRDLISRYELEKRYSEFSKOLEKALEE-----KNALLEN 961
Db 709 LVREKALALEVRLQA-VERDRODLAEOQLGSSAKELLESLFEAQOQNSVIEYVKGQLEV 767
QY 962 EYTCLSYK-FLPNEVECLAKNOISKASEEIMLQOEGEHSATISKQETIMQESQITLO 1020
Db 768 QIQTVQAKEVIOGEVRLCKLELDTL-----RSQAQOERDAARQ--LAQAQOEKTA 818
QY 1021 LTVDEVTHQSVQOOTEVOYLEMKHMDLFEKYIRNKSSEADLLREKMNLTGTMESVEVK 1080
Db 819 LEOOKAAHEKENVOLREKMEKERSMHO-----QELAKALPSLREKMELEMR 865
QY 1081 IADTKHELEETIRDKQOLHEKKYFFQAMQTIPTPLSDSLSPSKLVEGNSQDPIBND 1140
Db 866 LKEQOTEMEAIOAQREBERRQOAFSALCOMQ-----895
QY 1141 YHNILALATERNNIMVLETFERNLSLQOVIDNLQOLQASQESTKDLQPKDLEEGE 1200
Db 896 -----LETE---KERVSLLETLQTORELDAASQOULETRQDM---K 931
QY 1201 VKLLLEME---LKGHLTDSQLSIEKQLENLVETKQLTLOEMKNTTERNELOQNF 1257

```

```

Db 932 YOKIKEDETGIIQTOQEORRELEKEAARQH---RDILALOEESSSILLODKMDLOQOVE 988
QY 1258 DLAEHNSLKDODSENTEJOSTIODELRAAELEBQOKLOVDSFRDOLDCSVGISPNH 1317
Db 989 DIASQ--LVADDSORLVEO-EVOEKIRETOEYRIRIOKEL-----EREKASLSTLMEK 1039
QY 1318 DAVANDERVSIGEVNSLOE---MLGERDELOTSCRAVSELELIRAHVSVGENMLET 1374
Db 1040 E-----QRLVIVQEADSTRIOQELSLRQDMQEOEGEKELSAQEMELLNOVEKEKADPL- 1093
QY 1375 TKKLNGEKELEIKSESESEVYLKMLLENKDKNNKLEQAEYSEKEMQFSLIEVYFSSQK 1434
Db 1094 -----AOEALLELELSHTEQOL-----RASLMAOEAKAAQLOLRYSTES--- 1136
QY 1435 LVDEIEVLAQOLKAAAE-RLEIKRDYFELVOTANTNLVSGKLETPLQADH-----EE 1486
Db 1137 ---QLEALAEQOPGNAQOAOALASIALQALSGVCSRPRLSGGGSAPSVMGLER 1193
QY 1487 DSIDRSEEMEIVLGEKLEERNQYLLERLOEBKLEUS-NKL-ELLQEMETSVLLKDDLO 1544
Db 1194 DONGARS-----LKRGPPLTALSAAVAVASALHKLHODIMKTQOTRDVLRDQV- 1241
QY 1545 OKLESISENIIKENIDTFLKHSDTQAOLOKTQOELQAKNIALATASDNCPTQOKET 1604
Db 1242 OKLEERLTDIEAKSOYHTELO--DIOROLSONQEKSF-----WEGKON 1284
QY 1605 SADC-VHPIELEKILLLEELHOQKTNEQEKLLHKNELQAOVELKCEVHLKMSIERSK 1663
Db 1285 SLSELMELHETMASLOSRLRRAELQMEAGRELLQAKENILQAVEHLQAAVAVARA 1344
QY 1664 SLESLOHEKNDTEOQILALQOMQVYQEKKE---LOQTHEHLTAEDYHLEKNIETGLNF 1720
Db 1345 QASAGLTLEEDLTARSAKLKNEEVESEEREAQALOEOCELVYAQSKALQENIAL- 1400
QY 1721 KNEAQOKTTEKQCLNENKLEEOSQHLQCEIEELKSLDKESALETTESBQKVINLN 1780
Db 1401 -----LQOTLAREEVEETLRQOIQLEKQEMQKAALELLDLK---RN 1444
QY 1781 QEKEMVWLEMEIKNSORVIAERDLODLRESVEMSIETODDLKRAQALQOQKDYQ 1840
Db 1445 QEVDLOEQIOLEKFC-RSVL-----EHLRYAVQ-----ERBQKLVYORQIR 1486
QY 1841 EL-----TSQISVLOEKISILE-NOMLYNAVATKELISEDDLDNOSQNHLSFETILTSS 1894
Db 1487 ELKEKRETOKNVLEHOLELEKKQOMTE-----SQRGVQDLKQOLVT-LECLALE 1536
QY 1895 LKEKEFALE-----QAEKRDADARKTIDITEKISNIEQ-----LLOQATNL 1937
Db 1537 LEENHHKMEQOKLIKELBESQRTQORVALHHLTLEERSQELQAOSSQIHDELSHTVL 1596
QY 1938 KETLYERESLQCK---EOLANTHELRKTLKSKDLAKGMPQER---DEANKVYA 1988
Db 1597 ARBLQEDDOVKSQREQIELOKKEHLQODLERDQEL-MQOKERIQVLEDDRTQOTKI 1655
QY 1989 LFEKMSLSLEQINE---NVT-----LKESEGEKEFFYLRPSKQOSSQOMELRES 2037
Db 1656 LEEDDLQIKISTLERGRELTOORLOMORAEBSKG-----PSKAGQS-LEHKKLI 1705
QY 2038 LKTKDLOLEAEKEISEATNEIKNLTAKSISLEEILQNASILINENAVE---REMLRHS 2093
Db 1706 LRKREKEVEQOQHIEHLQELKQDLEQLOGLHRKRGESILLSQREQELVYLQOOLQRA 1765
QY 2094 KOOLVSELQOLSTLKSRRHAFASQSRKREDAVANKIASIAEELIKILTKEMDEFDRSKSL 2153
Db 1766 REQ--GELKEQSL-----QSOLDEAQRALAQORQOEALQOEQQAQNGQENR 1811
QY 2154 QBSSSHLSSEL---CTYKT---ELOMLKQOKEDINNNKLAEKYKEVDELLOHLSIKLEOL 2206
Db 1812 KKAADLOQALQEAQAHMTLERHGELODHKEQARRELEELAVEGRVY-----QALEEVL 1864
QY 2207 DQIOMELREKLNRYELCEKMDIMEKEISVLRMON--EPOQE-----EDDVAREMD 2256

```

```

Db 1865 GDILRAESR-----EOEKALLALQOQCAEQAEQEHVEVTRALQDSWLOQA 1908
QY 2257 ILSRNOEIOELMEKISAVYSEQHTLSSLSSELOKTEAKHCKMNLKISLSTLSRSF 2316
Db 1909 VIKERQDELALAE-----SOSKHOEBAANAR-----AEALDEALGKAI 1949
QY 2317 GSIQTEHVKLNLOTLINKFKVYRTAAVKEHSHILKDYEKDLAEQKRHDE---LR 2371
Db 1950 AALQG-----KEOHLLEQ-----AELSLSTASTATIQ 1977
QY 2372 LOQOCLQHGKRMKSDASSELKFCIEFLNELLFKKANIIQSVOODESEVQVNLQVGSF 2431
Db 1978 ASDDACQASHRQ---LEELRLIOEGEIDQD-----RYQEVQOIQALAQ----- 2021
QY 2432 LOEELERKKGFMQMLEFEGDLHADAKLSGMOQENRRIASTIQLTLKRLKAVAQSKIOR 2491
Db 2022 RDEELRHQERQEDLLEK-----SLAQVQENMTQOENQOGERE---EELRGLHQ--VR 2072
QY 2492 EITVYVNLQFPAKLOEKEQKELMRMEHNGPSASVMEENARLGLTKVODESKKLOS 2551
Db 2073 ELQTLAQKEQELIELEIRETOQRNNLEALPHSHKTSPEEBSLKLDSLEPRLQRELERLO 2132
QY 2552 RIKMLENE-----LNLVKDDAMHG--EKVALQDKLILSRNMAEELNANQVKL 2597
Db 2133 ALQOTEARELEMEKADODLALSLAQTKASVSLQEVAMFLQASVLEDSQO----- 2184
QY 2598 TKKODNLOAMKEIENLOKAVAKAVPYKE-----EIDMLKTVVAKIEMEKIR 2645
Db 2185 -RLQDELELTRALKER-ERHSHPGATSTBELSGRGQVGLQVSGVAEAPSPDGMKES 2242
QY 2646 YSKATQELAYLKSCLEDKEEGIRLKEELRRQADNDITVCYPKDYQKASTPVTYCGG 2705
Db 2243 W-----RQREHLOQAAVARLEIDRSRLQRRH-----VOLRSTLE---ODG 2279
QY 2706 SGIVQSTAMLV--LOSEKALEREIS-HYKKXYHHSRTWSSSEDRKRTAKSDAHSHT 2762
Db 2280 RQGRNSDQACVALLQEVYLLQALQILKQKODITYRKSQITRE-----L 2325
QY 2763 GSSHRSPPKTEYTRHGPVTPERSEMPSLHSGPKKSESSTKRVASPNSEIYSQLMSP 2822
Db 2326 AGLHSHLSHSLAVQAP-----EATVL-----EATRRL-----DESLOQLTSF 2366
QY 2823 GKTCMHKHLSPS 2835
Db 2367 GPVLLHP---SPS 2376

RESULT 6
US-09-820-843A-113
; Sequence 113, Application US/09820843A
; Publication No. US20030039963A1
; GENERAL INFORMATION:
; APPLICANT: Council of Scientific and Industrial Research
; TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE PRO
; TITLE OF INVENTION: USEFUL AS ANTI-INFECTIVES
; FILE REFERENCE: Q63915
; CURRENT APPLICATION NUMBER: US/09/820,843A
; CURRENT FILING DATE: 2001-03-30
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 113
; LENGTH: 2354
; TYPE: PRT
; ORGANISM: L. major
; FEATURE:
; NAME/KEY: misc. feature
; OTHER INFORMATION: AC005802_5 16202.3
; NAME/KEY: misc. feature
; OTHER INFORMATION: g116899670
US-09-820-843A-113

Query Match 5 6%; Score 821; DB 9; Length 2354;
Best Local Similarity 19.4%; Pred. No. 9e-24;

```

Matches 514; Conservative 474; Mismatches 1013; Indels 646; Gaps 91;

```

Qy 333 TPNVNEVDDALLKRYKKEIIDLK-----KOLENLESSSEFKKQAMKEHTOLLA--- 384
Db 123 SPQASVSTATLVHPVEDAVSTKPSVSEADLHAIIRITFELQALNDEGINALATLS 182
Qy 385 EIKOLHKREDRIWHLTIIVVASSQ-----ESQODRVKRRKRVTAAPGKIQNSLAS 437
Db 183 AAEOLRTAKKE-----NTAKSTAHLLQORLDITATQORAELEAVARLADRDRAAROL 236
Qy 438 GVSDFDLRLPGNFSKAKFSDMSPFEIDSVCTEESDDDALSMDSNGDAEMWLA 497
Db 237 AANNEELQORLDITATQORA-----ELEAVARLADRDRAAROLLAAN-----A 279
Qy 498 SKVHREKTSIHKSHMIDGQISDSQVHFDSKENGLOLVLPDSCGMACRKAPEKETS 557
Db 280 EELQORLDITATQOR-----AAELEAVARLADRDRAAROLLAANEE--- 320
Qy 558 LQOOLQSKKEEKELVOSFELKIA-----ELEOLSVKAKNLEMTNSREHSINAEVQ 610
Db 321 LQORLDITATQORAELE-----EAVARLADRDRAAROLLAANEELOQRDLTA---TQOR 371
Qy 611 TDVEKVVVRKMSVIGDSGYAASND--LQDSYDGRKLSSHDECIEHRKMLEOKIYDL 668
Db 372 AELEAQLARL-----ADRDRAAROLLAANEELOQRDLTAQORAELEAVARL 420
Qy 669 EEFLENLKSKSENDKSSBEDPMESIOLCEAIMAKANALEELALMDNDNTILEMET 728
Db 421 AANNEELQORLDITATQORAELEAVAR-----RLADRDRAAROLLAANEELOQRDLT 472
Qy 729 LKREIADLE-----RSLKENQETNEFEILEKETOKHEHAOLTHEIGSLKIVENAMQYN 784
Db 473 ATQORAELEAVARLADRDRAARQ-----QLAANEE-----E 504
Qy 785 LEEBLETKTKLLKEOFIOLAEIKKRAADNLOKKRVNFDLSVSKMGSEKICEEIPOLKOSLS 844
Db 505 LQORLDITATQORAELEAVARLADRDRAARQ-----QLAANEE-----QRLD 538
Qy 845 DAEAVTRDQACEGSLRESENELEKEMEDTSNMYNOKKKAASLEPEKOLETEKSNYKKNMEA 904
Db 539 TATQORAELEAVARLADRDRAAROLLAANEELOQRDLTAQORAELEAVARLADRDRAAR 598
Qy 905 DLQOELQSAFNEINYLNGLAG-KVPRDLRSVLELEKVFSEFSKOLEKALEKNALENEV 963
Db 599 ELQORLDITATQORAELEAVARLADRDRAARQOLLAANEELOQRDLTAQORAELEAV 657
Qy 964 TGLS-----EKFLEPNEVECKLNOISKASEEIMLIKGEHSASIIISKOETIIMQOSQOI 1018
Db 658 ARLADRDRAAROLLAANEELOQRDLTAQORAELEAVARLADRDRAAROLLAANEELE 717
Qy 1019 LQLDDEVYHTQSKVQOOTEEOYLEKMKHMDLFEKYIRKKSAPEDLREMEMLKGTMESEVE 1078
Db 718 QORLDITAT-----QORAELEAVARLADRDRAAROLLAANEELOQRDLTAQORAELE 771
Qy 1079 VKIADTHELEETIRDKQOLHKRYKFFQAMQIIFPIPLSDSLPPSKLVGENSDQPIEI 1138
Db 772 AQLARLADRDRAARQOLLAANEELOQRDLTA-----TQORAELE 809
Qy 1139 NDYINLALATERNNINWCVLETENRSLKEQYIDINTLOLOSIOAOSIEKSDLOKRODLEE 1198
Db 810 E-AQVAVARLADRDRAAROLLAANEELOQRDLTAQORAELEAVARLADRDRAAR 861
Qy 1199 GEVLLLEMLLEKHLTDSQSIEKLOENLEVEYEKLOTLQOEKKNITIEENLOTNED 1258
Db 862 ARQOLLAANEELOQRDLTAQORAELEAVARLADRDRAAROLLAANEELOQRDLTAQORAELEAVAR 921
Qy 1259 LKAHDSLKODLSENIOSIETODELRAAORELEKOLVDSFRQOLDSVSGISSPNHD 1318
Db 922 LAADRDEARQOLLAANEE-----LQORRD----- 945
Qy 1319 AVANQEVYSLGEVNSLOSSEMRGERDELQTSCKALVSELELLRAHVSVSEBENLEITKLL 1378
Db 946 -TATQORAELE-----EAQLARLADRDRAAROLLAANEELEIQ-----QRLDTA--- 986

```

```

Qy 1379 NGLEKEIIGKSESEVLSKMLENLKEDNNKLEQO-----AEYSSKENQFSLEEVSSGSK 1434
Db 987 -----TQORAELEAVARLADRDRAAROLLAANEELE-----QO 1020
Qy 1435 LVDEIEVLKAOQLKAAEERLEIKRDYFELVQVANNLEVEKLEPPLQADHEEDSIDRSSE 1494
Db 1021 RLDITATQORAELEAVARLADRDRAAROLLAANEELOQRDLTAQORAELEAVARLADRDRAAR 1069
Qy 1495 EMEITVYLBKELERNQYLLERLOEBKLELSNKLIEITQEKETSVLLKDDLOKLEBSLESEN 1554
Db 1070 ELEAVARLADRDRAARQOLLAANEELOQRDLTAQORAELEAVARLADRDRAAR 1121
Qy 1555 IIKENIDIT---LKHSDTQOQLOKTOOEIOLAKNLAIASDNCPIQEKETSDCVHP 1611
Db 1122 DEARQOLLAANEELOQRDLT-----ATQORAELEAVARLADRDRAAROLLAANEE---E 1173
Qy 1612 LEEKILLTEELHOKTNQOEKLLHEKNELEQAOVELKEVEHLMKSMIESKSSLES----- 1667
Db 1174 LQORLDITATQORAELEAVARLADRDRAARQOLLAANEELOQRDLTAQORAELEAVAR 1233
Qy 1668 LQHEKHTEBOQLA-----LKOQOYVTOEKKELOOTHEHLTAENDHLKENTELGLNFKNE 1723
Db 1234 LAADDEARQOLLAANEELOQRDLTAQORAELEAVARLADRDRAARQOLLAANEELE 1291
Qy 1724 AOKTKTEQCLNENKELQOSQHRLOCEIEELMKSL-----KOKESA 1765
Db 1292 QORLDITATQO-----QORAELEAVARLADRDRAAROLLAANEELOQRDLTAQORAELEAVARLADRDRA 1347
Qy 1766 LETTESQKVYNLQOEMKXVLEMEELKNSQRTVYARSDLOQDLDRSVEKSIETODDL 1825
Db 1348 ROQLAANEE-----LQORLDITATQORAELEAVARLADRDRAARQOLLAANEE---ELQORL 1401
Qy 1826 RKAQALQOQKDKVQELTISOISVLOEKISLENOVLNVAIVYKELTSLERDLDNGSKOHLF 1885
Db 1402 DTATQORAELEAVARLADRDRAARQOLLAANEELOQRDLTAQORAELEAVARLADRDRAAR 1458
Qy 1886 SEIETLSLKEKEFALQAEKDKADARKTTIDIREK-----ISNIEE---QLQOATNLKE 1939
Db 1459 ANAELEQORL-----DTATQORAELEAVARLADRDRAARQOLLAANEELOQRDLTAQORA 1515
Qy 1940 TLYERESLI-----QCKBOLALNTEHLEETLSK-----DLAOKMEORDEANKVY 1987
Db 1516 ELEAVARLADRDRAARQOLLAANEELOQRDLTAQORAELEAVARLADRDRAAR 1570
Qy 1988 ALTEKMSLEQINENYTTLKEGSEKETFYIQ-----RPSKQOSSQOMEELRESEKTKRD 2042
Db 1571 --ROQLAANEELOQRDLTAQORAELEAVARLADRDRAARQOLLAANEELOQRDLT- 1626
Qy 2043 LQLEBAKEKISPATNEIKNTAKISLEEEIILQNASIINEAVSERENRHSKQVLELE 2102
Db 1627 -----ATQORAELEAVARL-----AADDEAR---QOLLAANEE 1657
Qy 2103 QLSITLKSRDHAFAQSKREKDEAVNKIASLAEIKILTKEMDEFDRKESIOEGSSHLSSE 2162
Db 1658 ELQORL-----DTATQORAELEAVARLADRDRAARQOLLAANEE---QOLLAANEE 1696
Qy 2163 ELCTYKTELOMLKQOKEKEDINKKLEKVEVDELLQHLSSLEQOLDQIOMELNENLRYNE 2222
Db 1697 EL---QORLDITATQORAELEAVARLADRDRAARQOLLAANEELE---QORLDITATQORAE 1750
Qy 2223 LCEKMDIMEKETSYLR-----LMQNEPQOGEEDVAVARMQDLSRNOEIOELMKKI 2272
Db 1751 LEVEAVYLRRERREARGETAVAGBEVOLYRTEVEEELCKEERKCLSESRVQOLR----A 1806
Qy 2273 SAVYSEQHTLLSLSELQKETE-AHKHCLMLIKESISLTSRSFGSLQTEHVKLNTOLO 2331
Db 1807 SAAKQOQROEVAAKANEVQERDLSMARCIATHGGAPOGADGRDQALQOLANLREEVYLS 1866
Qy 2332 TLINKFKVYVYRT-AAVKEDSHLIDYEDKDLAEOGRHDELQLOCLQOGRKSSDSASE 2390
Db 1867 ---EKOKAMERYIPQVRRORMLR-----AAEBORAD---LEARLVDEADGLDRSPAAVS 1914

```



```

Db 1178 -----ELQORL---DTATQORAELEAVARLANAAEELQORLDTATQORAELEQARLA 1230
Qy 1670 HEKHDTBOQLA-----LKQOVVTOEKKELQOTHEHLTAVDHUKENIELGLNKEAQ 1755
Db 1231 ADREARQOLAAAEELQORLDTATQORAELEAVARLANAAEELQORLDT-----TATQ 1284
Qy 1726 OKTKBEOCLINENKELBOSQRLQCEETELMKSLKDKESALEETLKESQKYNINQEMEM 1785
Db 1285 QRALEAVARLANAADRDEARQOLAAAEELQORL---DTATQORAELEAVARL----- 1335
Qy 1786 VMEEMELKNSQRTVIAERDQLQDRLRESYEMSIETODDLKAOBALQOQKDYQELTSQ 1845
Db 1336 -----NADGEARQOLAAANE-----ELQORLDTATQORAELEAVARLAN 1377
Qy 1846 ISVLQERTISLENOMLYNATVKTLSERDDLNSKQHLSEITTLISLKEKFPALBOA 1905
Db 1378 AEEELQORLDTATQORAELEAVARLANADR-----EARQOLAAAEELQORL---DTATQOR 1431
Qy 1906 EKDKADARKTIDITEK-----ISNIEE---OLLQOATNLKERTLYERESLI-----QCKEOL 1954
Db 1432 AELEAVARLANAADRDEARQOLAAAEELQORLDTATQORAELEAVARLANAADRGEARQOL 1431
Qy 1955 ALNTEHLRETLKSK-----DLALGKWEDEBANKVIALTEKMSLEBQINENYTTL 2007
Db 1492 AANAEELQORLDTATQORAELEAVARLANAADRDEA-----RQOLAAAEELQORLDTA 1544
Qy 2008 KEGGEKETFLQ-----RPSKOSSSOMELRESLTKDLQLEAEKEISEATNEIKNL 2052
Db 1545 TQORAELEAVARLANAADRGEARQOLAAAEELQORLDT-----ATQORAE 1590
Qy 2063 TAKISSLEEETILQNASILNEAVSERENLRHSKOQYSELBOLSTLKSRRHAFQSKREK 2122
Db 1591 EAVARL-----AADREAR---QOLAAAEELQORL----- 1619
Qy 2123 DEAVNKIASIAEETIKITKEMDEFDSKESLQEOSSHSELCTYKTELQMLKOQKEDIN 2182
Db 1620 DTATQORAELEAVARLANAADRDEAR-----QOLAAAEEL---QORLDTATQORAELE 1669
Qy 2183 NKLAEKKEVDELQHLSSLKEOLDQOMELRNKLNVELCEMDIMEKISTVLR----- 2238
Db 1670 AQLARLANAADRDEARQOLAAAEEL---QORLDTATQORAELEAVARLANREARBEA 1726
Qy 2239 -----LMONEPOOEDDVAERMIDLSERNOIEQELMEKISAVSEOTHTLSSELOK 2292
Db 1727 VAGQOVLYETVEEBEELCEKRWCLSRVAQLRE-----ASAAKQORQOEAANAANEOE 1782
Qy 2293 ETEA-HKCHMLNKESLSTLSRSFSGSLQTEHVKLNTQLOTLNKFVYVYRT-AAVVEDH 2350
Db 1783 RLDSMARRCIAHEGDAPORADGRDALRQLANLREEVKLS---EKOKAMERVIPGVBRO 1839
Qy 2351 SLIKDYEKDLAAEQKRHDELRLQJOCLEQHGKRSKSDASSELKCELEFNLLEFKKANI 2410
Db 1840 MRLE-----AAEQRAD---LEARLYDEAGDLRSRPAST---NEVNLVRLDLQEH-- 1885
Qy 2411 IOSVODDFSEVOYVNLQVSTLOELEHKKGFMQMLEEFGDLHVDAKKLSGMOENRRI 2470
Db 1886 -EAAQNNCTTLEA---QVASLTSR-----DNGRQOESGDL 1917
Qy 2471 ASTIQLTLKRLKAVOSKIOREITVYNQFEAKIQEKKKONKELMRMHEHGPSASVME 2530
Db 1918 -----SEAOHRHLD-----NVOE-----ROMAHH--RCALAE 1942
Qy 2531 ENARLGLITLTVODESKKLSRIKLENEL-----NLYKNDAMHK- 2570
Db 1943 QNAAASELQAVKAKLKQASVYKASSLMTRLSASSGAGCVSARVYVGGSSAVPOAAHRD 2002
Qy 2571 GEKVALIDKILSRNAEELNAMOVLTKKODNLQAAKMEITENQKVVAGAVPYKEID 2630
Db 2003 AELIAEGERLBERGEMARLLAEGVELRERARPLEVLAE-----KLIGDRSTDAEVA 2057
Qy 2631 NLKTKVYKIEKIKYKADQETAYL---KSCLEDEKEBGRILKEELR----- 2676
Db 2058 TEPTOV---RRNAASHRHLDSREAOIDERAARLREKEOQLLRVARELQTSRALQVLYA 2113

```

```

Qy 2677 -----RAQ-----ADNDTYC---VPKDYQKASTFPVTCGGSGIVOSTAMLVLOSEKA 2723
Db 2114 RALNRPOVTSLLTADGDISTYPPDTPOOQOQTRTP-----LREPVYS 2156
Qy 2724 LERELSHYKKKYHHLSTWSSSEDRKTKAKSDAHSHTGSSHRGSPKHTETYNHGPVP 2783
Db 2157 LDSEYAH-----GRTAGAAV-----SSGLASPL-----PREP 2184
Qy 2784 ERSEMPSLHLSPPKSSSTKRVVSPN---RSEITYSOLVMSPGKTM 2827
Db 2185 PRARVYHRAVATGTEDTYVRLTAATEARYADVLYEHLETSNGLOGV 2231

RESULT 8
US-09-820-843A-114
; Sequence 114, Application US/09820843A
; Publication No. US20030039963A1
; GENERAL INFORMATION:
; APPLICANT: Council of Scientific and Industrial Research
; TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE PRO
; FILE REFERENCE: 063915
; CURRENT FILING DATE: 2001-03-30
; CURRENT APPLICATION NUMBER: US/09/820, 843A
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 114
; LENGTH: 2310
; TYPE: PR
; ORGANISM: L. major
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: AC005893_12 L6202.3
; NAME/KEY: misc_feature
; OTHER INFORMATION: g1|6899664
US-09-820-843A-114

Query Match 5.58; Score 811; DB 9; Length 2310;
Best Local Similarity 19.48; Pred. No. 2,1e-23;
Matches 475; Conservative 450; Mismatches 926; Indels 596; Gaps 84;

Qy 558 LQOQLOSKSEKKELVQSFELKIA-----ELEEOISVYKKNLEMYNRSRHSINAEVQ 610
Db 204 LQORLDTATQORAELEAVARLANAADRDEARQOLAAAEELQORLDTA-----TQOR 254
Qy 611 TDVEKEVYRKEMSVLQSGYNASNSD---LQDSYVQGRKLSSSHDECIEHRRMLEOKIVDL 668
Db 255 AELBARVARL-----AADREARQOLAAAEELQORLDTATQORAELEAVARL 303
Qy 669 EEFTEINLKKSEKENDKQSSBEDFMESTQLCRAIYAKKANALEETAL---MRDFNDITL 724
Db 304 AANAEELQORLDTATQORAELEAVARLANAADRDEARQOLAAAEELQORLDTATQ 358
Qy 725 ENETLKREIADLSERKNOETNEFEILEKETQKEHEQOLIHEIGSL-----KKLVE 776
Db 359 QRALEAVARLANAADRDEARQOLAAAEELQORLDTATQORAELEAVARLANAADRDEARQOLAA 411
Qy 777 NAEYNTONLEEDLTKTKLKEDEIQLAELKRRADNLQKKVVRNFDLSVSMGDSKLCCEI 836
Db 412 NAE-----ELQORLDTATQORAELEAVARLANAADRDEAR-----QOLANA 452
Qy 837 FOLKQSLSDAEAVTRDQKQESFRLSRNLELEKEMEDTSNMVYNQKKAASLFEKQLETEK 896
Db 453 EELQORLDTATQORAELEAVARLANAADRDEARQOLAAAEELQORLDTATQORAELEAVAR 512
Qy 897 SNYKMEADLOKELOSAFNEINLYNGLAGKVPBDLSRYELEKVS-----EFSKOL 949
Db 513 ARLAADRDEARQOLAAAEELQORLDTATQORAELEAVARLANAADRDEARQOL 565
Qy 950 EKALEEKNALDENETCYSEYFLDPNEVECLKNQISKASEIMLLKQGEHSASTISQET 1009
Db 566 DTATQORAELEAVARLANAADRDEARQOLAAAEELQORLDTATQORAELEAVARLANAADRDEARQ 619

```

QY 1010 IMOSEBOLITDEVTHTOSKVOOTEBOYLEMKHMDLFEKYIRNKSEADILLREMEN 1069
 Db 620 QLANAEELQRLDTAT-----QORAELEQRLAADGEARQOLAAAEELQRLDT 673
 QY 1070 LKGTMESVAKIADTKHELEETIRDKOLLHEKKYFFOAMOTIPITPLSDSLPSKIVE 1129
 Db 674 ATQORAELEQRLAADGEARQOLAAAEELQRLDTA----- 713
 QY 1130 GNSODPIENDYHNLALATERNNIMVCLTERNSLKEQVYIDLNTOLQISQASIE-KSD 1188
 Db 714 --TOQRELE--AOLARLAADGEARQOLAAAEELQRLDTATQORAELEAVARLAAD 769
 QY 1189 LQKPRDOLBEREYVLLLEMLKGLTDSQSLKLTOLLEMLVEYKLT-----QTLQEMKN 1244
 Db 770 GDEARQOLAAAEELQRLDTATQORAELEAVARLAADGEARQOLAAAEELQRLDT 829
 QY 1245 ITTERNELQTFEDLKAEHDSLKODL---SENIEQSIETODELRAA-----QOE 1290
 Db 830 ATQORAELEQRLAADGEARQOLAAAEELQRLDTATQORAELEAVARLAADRE 889
 QY 1291 LBREKOLVDSFROOLLCSSGIVSSPNHDVAVANQKVSLSGEVNSLSQSMLENGERDELQTS 1350
 Db 890 ARQOLAAAEELQRLD-----TATQORAELE--EAQARLAADREARQOL 933
 QY 1351 KALVSELEL-----LRAHKSVEGENLEITKRL-----NGLEKEITLKSSESEVL 1395
 Db 934 AANAEELQRLDTATQORAELEAVARLAADGEARQOLAAAEELQRLDTATQORAELE 993
 QY 1396 KSMLENKKNKKLKEQ---AEYSKKNQSLSEVFSQSLKLTVEYVLAQOLKAAEE 1451
 Db 994 EAQARLAADREARQOLAAAEEL-----QORLDTATQORAELEAVARLAAD 1037
 QY 1452 RLKIDRYELVOTANTNVEGKLETPLOADHEEDSIDRSEEMETKVLGKELERNQYL 1511
 Db 1038 RL-----AANAEELQRLDTATQ-----QRAELAVARLAADRE-- 1072
 QY 1512 LERQOEKLELSKLETLQEMETSVLLKDDLOKLESLSENITIKENITDTLKHSHDT 1571
 Db 1073 -----EAKQOLAAAEELQRLDTATQORAELEAVARLAADGEARQOLAAAEELQRLDTATQORAELE 1127
 QY 1572 QAOLQK-----TOELOLAKNLALIASDNCPIQOEKTSKDCVHPLEKI 1616
 Db 1128 EAQARLAADREARQOLAAAEELQRLDTATQORAELEAVARLAADREARQOLAAAEELQRLDTATQORAELE 1177
 QY 1617 LITTEILHQTNEQKLEHEKNELEQAUYELKCEVEHMKSM-----TESKSSLESIO 1669
 Db 1178 -----EELQRL-----DRAQORAELEAVARLAADGEARQOLAAAEELQRLDTATQORAELEAVARLAAD 1230
 QY 1670 HEKHDFEOLLA-----LKQOMVYVTOEKKELOOTHEHLTADEVHLENLELGINFKNQAO 1725
 Db 1231 ADDGEARQOLAAAEELQRLDTATQORAELEAVARLAADGEARQOLAAAEELQRLDTATQORAELEAVARLAAD 1284
 QY 1726 OKTKTEOCLINENKLEESQHRLOCEIEELMKSLKDESALFETIKSEOKVINLQEMEM 1785
 Db 1285 QRAELAVARLAADREARQOLAAAEELQRL-----DRAQORAELEAVARLAAD 1335
 QY 1786 VMELEELKNSQRTVIAERDQLODDIRESVENSIEFTODILRKAQOALQOKDKVOELTSQ 1845
 Db 1336 -----AADGEARQOLAAAEEL-----ELQORLDTATQORAELEAVARLAAD 1377
 QY 1846 ISVLOEKISLENQMLYNATVYKETSERDLDNOSKOHFSELETLSLKEKEFPLBOA 1905
 Db 1378 AEELQRLDTATQORAELEAVARLAADRE--EAKQOLAAAEELQRL-----DRAQOR 1431
 QY 1906 EKDKADARKTIDITEK-----ISNIEE--OLLOQATNLKETTYERESLI-----QCKEOL 1954
 Db 1432 AELEAVARLAADREARQOLAAAEELQRLDTATQORAELEAVARLAADGEARQOL 1491
 QY 1955 ALNTEHLRETLKSK-----DLALGMEORDEANAKVIALTEKNSLSSEQITENENTVL 2007
 Db 1492 AANAEELQRLDTATQORAELEAVARLAADRE-----ROQOLAAAEELQRLDTA 1544

QY 2008 KEGEKEFETYLQ-----RPSKOSSSQWELRESLTKTDLOLEAKELISEATNEIKNL 2062
 Db 1545 TOQRAELEAVARLAADGEARQOLAAAEELQRLDT-----ATQORAELE 1590
 QY 2063 TARISSLEELIIONASILNVAERENLRHSKOOLVSELEQSLTLKSRDHAFQSKREK 2122
 Db 1591 EAVARLAADREARQOLAAAEEL-----AADREARQOLAAAEELQRL----- 1619
 QY 2123 DEAVNKIASIALEIKILITKEMDEFDSKESLQSSHLSSELTCTYTELQMKQKEDIN 2182
 Db 1620 DRAQORAELEAVARLAADREAR-----QOLAAAEEL-----QORLDTATQORAELE 1669
 QY 2183 NKLAKEYEVEDDELLOHLSLKEQLOQLOMELREKRLNLECEKMDIMEKETSYLE----- 2238
 Db 1670 AQARLAADGEARQOLAAAEEL-----QORLDTATQORAELEAVARLAADREARQOL 1726
 QY 2239 -----LMONEPOEEDDVAERMDLESNRQETLOELMERISAVYSQHTLSSLSSELOK 2292
 Db 1727 VAGEQVQVLYREVEEBCLEKEERWCLSRVQALRE-----ASNAKQORQOEVAKANEVOE 1782
 QY 2293 ETEEA-KHKHMLNIKESLSSTLSFSGSLQTEHYKLTQLOTLNKKVYRT-AAVKEH 2350
 Db 1783 RUDSMARCTIAHGDAPOQADGHDALROLANLEEVKLS--EKQKAMERYIPGVERQ 1839
 QY 2351 SLIKDEKDLAABOKRHDELRLQLOCLEQHGKRWSDSASEELKFCIEFLELLEFKANI 2410
 Db 1840 MRLE-----AAEQRAD--LEARLYDAGDLSRPAAT--NEVNLRLDALQEH-- 1885
 QY 2411 IOSVODFSEVOVFLNQVOSTLOEELHKKGFQWQMLEFQDLHVDKLLSEGMQOENRI 2470
 Db 1886 -EAOQRCTTLEA--QVASLTSDR-----DNGROQESADL 1917
 QY 2471 ASTIQLTRKLAQVQSKLOREITVYVUNQFEAKLOEKKQONKELMRBHEHGPSASYME 2530
 Db 1918 -----SEAOHRHD-----INQD-----RDMANH--RCALAE 1942
 QY 2531 ENARLIGITVODESKRLQSLRIKLENL-----NLVKDDAMHK- 2570
 Db 1943 ONAAMASELOAVAKRLQASVYKASSIMTRLASASSGAGVSAVRVGSASVAPQAPHRD 2002
 QY 2571 GEKVALLQKLSRMAEALNMOYVLTQKQONLQAMKEIENLQKMAKGAVPKEED 2630
 Db 2003 AELIAVGERLREKRGAMRLABGVELREARPLEVYLE-----KLIDRRTSDAEVA 2057
 QY 2631 NLKTVVKKIEMEKIKYSKATQDEIAYL--KSCLEDKEGLRLKEELR----- 2676
 Db 2058 TEPTQV-----RRNAHSHRLDSREAOQDLDERAARLREKEQQLLRVARELOTKSRALQVLYA 2113
 QY 2677 -----RAQ-----ADNDTVC--VPKYQKASTFPYTCGGSGSIVYSTAMLYQSEKRA 2723
 Db 2114 RALNRQVYVSLTLADGDTSTYPTDPOQOQOQGTPT-----LREPYVS 2156
 QY 2724 LERELSHYKKYKHHLSRTUSSSEDRKRTAKSDAHSHTGSHRSPHKTETRYRGPYTP 2783
 Db 2157 LSEVAHY-----GRTGAAY-----SSGLASPL-----PREP 2184
 QY 2784 ESEEMPSHLGSPKRSSESTKRVVSPN--RSEIYSOLVMPSPKTCM 2827
 Db 2185 PRARVHRAVEATGTEEDTQVRLTAATEAYRVLYEHILLESNGLGCV 2231

RESULT 9
 US-09-874-923-120
 ; Sequence 120, Application US/09874923
 ; Patent No. US20020081320A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Reed, Steven G.
 ; APPLICANT: Campos-Neto, Antonio
 ; APPLICANT: Webb, John R.
 ; APPLICANT: Dillon, Davin C.
 ; APPLICANT: Skeiky, Yasir A.W.
 ; APPLICANT: Bhatia, Ajay
 ; APPLICANT: Colter, Rhea


```

QY 2351 SLIKDYENKLAAROKHDELRLQLOCLBQGRKWSDSASELKFCEIEFLNELFFKANI 2410
Db 1840 MRLE-----AAEQRAD-----LEARLVDEAGDLRSRPAAT-----NEVNLRYDLALQDEH-- 1885
QY 2411 IOSVDQDFSEVQVFLNQVSTLOEELEHKKGFQMWLEFGDLHVDAKTLSEGMQENRRI 2470
Db 1886 -EAGQRCITLLEA---QVASLTSR-----DNGQQSASDL 1917
QY 2471 ASTIQLLTRKLVAVOSKQIQREITTYVLYNOFEAKLOEKKOEKNEELMRMEHGFASVMBE 2530
Db 1918 -----SEARHLID-----NVQE-----RDMANH--RCALTEE 1942
QY 2531 ENARLILGILKTVDQESKQIQRIMKLENEL-----NLVKDAMHK- 2570
Db 1943 QNAAASELOAVAKLRQASVYKASSLMTRLSASSGAGVSAVRVGGSSAVYQAAPHD 2002
QY 2571 GERVAILQDLKLSRNAEAEINAMQVLTTRKODNIQAAMEIENIQAQVAKGAVPYKEED 2630
Db 2003 AELIAVEGERLBERGEMRLLAGVELREBARPLEVLAE-----KLIGDRITSDEAVA 2057
QY 2631 NKTQVVKIEMEKIKYSKATDQIALV---KSLDEKEGRLKLEELR----- 2676
Db 2058 TEPTQV---RNAASRHLDSEKQDLDERAARLREKEQQLKRVARELQTKSRALQVLYA 2113
QY 2677 ---RAQ-----ADNDTVVC---VPRDYQKASTFPVTCGGSGGIQSTAMLVLOSEKAA 2723
Db 2114 BALNRQVAVSLTLTADGDTISYPTDPOQOQGRTP-----LREPVYS 2156
QY 2724 LERELSHYKKYHHLSRTMSSSEDRKRTAKSDAHSHTGSSHRGSPHTEYRHGPVMP 2783
Db 2157 LPSEVAHY---GRTAGAAV-----SSGLASPL-----PRBP 2184
QY 2784 ERSEMPSLHSGSPKSESSTKRVSPN---RSEIYQOLVWSPGKTGM 2827
Db 2185 PRARMVHRAVEATGTEEDTQVRLTAATEAVRDVLYEHILLESNGLOV 2231

RESULT 10
US-10-023-219-4
; Sequence 4, Application US/10023219
; Publication No. US20030032592A1
; GENERAL INFORMATION:
; APPLICANT: Myriad Genetics, Inc.
; APPLICANT: Cimboira, Daniel M.
; APPLICANT: Heichman, Karen
; APPLICANT: Bartel, Paul L.
; TITLE OF INVENTION: Protein-Protein Interactions
; FILE REFERENCE: 2318-282-II
; CURRENT APPLICATION NUMBER: US/10/023,219
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 60/256,983
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 2139
; TYPE: PRP
; ORGANISM: Homo sapiens
US-10-023-219-4

Query Match 5.4%; Score 794; DB 9; Length 2139;
Best Local Similarity 20.9%; Pred. No. 8.7e-23;
Matches 527; Conservative 444; Mismatches 912; Indels 634; Gaps 113;

```

```

QY 364 -----ESSETKAQAMAKEEH--TQLLAETIKLHEREDRIWHLNTIVASQSES 411
Db 127 PEVYVIEPIDEEARPSHITPAGDSEHMTQ---RSEIYAEQQLKRWNDLINAQSGSS 183
QY 412 -QODQVRKRRRYVAPAGKISLHSAVSPDMLSLPGNSKAKKESDMPSPEDIDS 470
Db 184 PRPD-----WIEKIQVECEDIGITRDGHLNR-----KLVLSICEQGLQNVQDE 228
QY 471 VCTEF--SDPDALSMMD-----SNGIDAEWNLASKVTREKTSIHQSMDFQG---- 517
Db 229 MLEEVFNHLDPOGTMSEVDFPYGLFKNGSL---TPSASTPYQLKRLHLSMQSDESGCR 285
QY 518 -----ISDVQFHDSKKNOLVYLPKDSGDMAECKSAFEKEITSLOQO--LOSKEEERK 570
Db 286 TITSSAMTSTIGR-----VFSCLDDG---NCHASVETHLDTWQEGEIGNSGELK 333
QY 571 ELYQSEELKIALEBOLSYKANKLEVYNSREH-----SINAEVQTDVK--EYVRKES 623
Db 334 ALDFSIDGNINILTE--LTTLARENELVYTKNSIHOAALAFKAEIRHLRLERVOVVAEKER 391
QY 624 VIGDSGYNMSNDLODSSVDGKRLSSHDECIEHRKMLQKIYDLE--EFIEHLNKKSEN 681
Db 392 LNSDLDAEKAKLSLASEVD-----DHAAIFERRNEYVLRKLDGEYKRIALAKNELRK 445
QY 682 DKQSSQEDFMESIOLCEAIIAEKANALBELALMRDNFDNILLNETLKEIADLERSLK 741
Db 446 ERQOIIQAQKQKQLELEQEI--EKAKTEEN--YIRDR-----TALSILK 484
QY 742 ENQETNEFLEKEIENKHEADLIHEIGSLKLVENAEVYNNONLEEDLETYKRLKEQEI 801
Db 485 ENSR-----LENE-----LENAE----- 498
QY 802 QLAELKRRADNLQKVRNFDLSVSMGDSSEKICEEIFOLKQSLSDAEAVYTRDAQKECSFLR 861
Db 499 KLAIEYNLNLKQIRNLNEN--LAEKFGDLDPSAEEFLQGEERLQO---MNEYERQCRVLQ 554
QY 862 SENLEKEMEDTSNMYNOKERKASLFEKQLETEKSNYKMEAD---LOKELQSAFNETN 918
Db 555 DOYDELQSELEE---YRAGRYL---RLPLKNSPESEVENSGGIRPEGLGSEECN 605
QY 919 YNLGLLAGKV-----PRDLTS--RVELEKKVSEFSQOLKALBE--KNALEN--EVTCL 966
Db 606 PLMNSTEAEVLVEQMEQHNRDCCRLLEDEKVRITYEQLDBETVYSCKAQENMKQRIE 665
QY 967 SEYKFLPNEVECLKNQISKASEIMLKQEGHSASISIKOETIMQOSEQIQLTDEVY 1026
Db 666 NETRIEKLQISDLKNIAELQGAVALK--EAMHEAV--CRHE---BEKQLOVYLEEKT 719
QY 1027 HTQSKVQOTEBOYLEKKKHHDLFEKYTRKNSAEADLLFEMENLAKTMSVSE-----VK 1080
Db 720 HLOEKRLQHE--MELKARLTQAQASFERERGLQSSAVTEERKVRGLTLELQFHOEQLT 777
QY 1081 IADTKHELEETIRDKQOLHEKKYFQAMQTLFPIPLSDSLPPSKLVGNSQDPLEIN- 1139
Db 778 SLVEKHTLE-----KEEL--KRELEKKQ-----RLQGRKREKMEIECRH 815
QY 1140 -----DYHNLIATATERNNIMVC---LETERNSLKEQVIDLNTQLOSLAQSIKESDL 1189
Db 816 RISOIAPQFQSDQKVTYR-----CESALQSLGGRYRQELKDLQEQ-----QREKSSQM 864
QY 1190 QRPKQDL--EEGKVILLLEMLLKGHLTDSQSLTEKLOLENLEVTEKQLOTOEMKNITI 1247
Db 865 EEEKDLQTECAEAOBLLETKRKTTSVLTQEREME-----KTYEHLNLSMVV 916
QY 1248 ERNELOTFNEDAKEDHSIKODLSENIEQSIETQDLRAA--QEBLREKQLOVDSFRQULL 1306
Db 917 ERQOILLQIEDLRNVSEITQOSLSDQIL-----ELKSSHKRELRERREVL----- 961
QY 1307 DCSVGISSP-----NHDAVANOEKVS---LGEVNSIQSEMLGEBDELOTQSK 1351
Db 962 -CQAGASQELQASRLERLMEHQ--ERQEMMSKLLAMENIHKATCETADRAEMNST--- 1016
QY 1352 ALVSELELLRAHVSVGENLEITTKLNGLEK--ELIGNSE--ESEVLKSMI---ENLAKED 1405

```

```

Db 1017 -----EISRLQSKIKEMQ-----QATSPILMLOSGOVYIGEVEBEDGALSTILOOGEOLLEE 1068
Qy 1406 NMLK-----KEBAEYSSKKNPFLSEYVSGSKLVDEIVLKAQJKAABEELKID----- 1457
Db 1069 NGVLLSLQRAHBOAVKEN-----VKMAEISRLQOOLQKLEPIVSSCLDEP 1117
Qy 1458 -RDYF-----ELVOTANTNLVEGKETPLADHEED-----SIDRS---E 1494
Db 1118 ATEFFGTAQOTQOFLQONTKQOEVTRRHVLSLEDEEDVRLGSTGTSVQROEKIE 1177
Qy 1495 EMEIKVYG-EKLEBNOYLLERLOEKELESNKLEILOKEMETSVLLKDDILOKLESILSE 1553
Db 1178 ESASVEGFSELENS-----EETRTSEWELKNQISQLOEQMLMCAODRASSEKKODLLFD 1233
Qy 1554 NILKKEI-----DTTLKH-----HSDPTQAOLOKQOEOLOLANLAAADNCPITQOKE 1603
Db 1234 VSVLKRLKMLKLERIPKSPRYKLLYEDVSRENDCLOELMRMERYDEALEN-----NKE 1288
Qy 1604 TSADCVHPLKEKILLLEELHOKTNEQEKILHEKNELEQAQVE-----LKCEVH 1653
Db 1289 LTAH-----VRLQDELKMKMEVYETPLSEKSYDEKVIENEGNLVYLRLQKIEK 1340
Qy 1654 LKSMIES-----KSSLESLOHEKHDTEOQLLALQAOQOVYQEKKELOQTHEHILTAEV 1707
Db 1341 LQESVQORCCCLMEASLEWLEIE---PDGNILQNLQTEBECYPRVSRHVHVEECKOEN 1397
Qy 1708 DHLKENELGLKNKNEQOQTTKEQCLLNKNEKELOEQSHLOCEI---EE--LMSKLKDE 1763
Db 1398 QYLEGNQUL-----LEKKAHEIAWLGITQTHORPRVQOVYVILLEETTLTGFDCK- 1449
Qy 1764 SALETLESEKQVYNLQOEWMLMEELKNSORVIAERDQLODDLRSSEKISITOP 1823
Db 1450 -----HQHQQTIAELELEKTK-----LOELTRIKERVITVLQOKD 1486
Qy 1824 DLKKAQALQOQKQVQELTSQISVLQEKISL-----ENOMLVNAYVQKELTSE 1874
Db 1487 VLSHGEFE-BELKAMMHDLQITCEMOQKVELLRYESEKLOEQENSILRNET--TLNEE 1542
Qy 1875 DDLNQSKOHFLSEIFETLSLSLKEKEFALBEQEKDKADMAKKTIDITKISINIEQOLQA 1934
Db 1543 DSIISNKK-----LGTINSQOEEWOMKTEVTKOENNAVQKVELKQOIS--ELKIKINOQ 1594
Qy 1935 TNLKET-LYERESLQCKEOLALNTEHRLTKSKDALCKMBOERDEANKVYALTEKM 1993
Db 1595 LDEPNTLESKNS--QOQEKLOELNQLTLEMOKEKEPKN----- 1633
Qy 1994 SSLEQOINENVTTLKEGEGEKETFYLOP--SKOQSSQMEELRESLTKTKDLOLEBAEK 2050
Db 1634 SALEER-----BOEKNLKEELERCKQVQSSSTLYSLEALSEVKIQTHIVQO 1680
Qy 2051 E---ISEATNEIKNL--TAKISSLEBEILONASLINEAVSERENLRISKOOLYSELOLS 2105
Db 1681 ENHLLDELLEKMQHLRCPDISFOOKISSVLSYNEKLKEKALSBEELNSCVDKLAKSS 1740
Qy 2106 LTLKSRDHAFAQSKREDEAVNKRISLAEIEKILITKEMDEFRDSKESLOSSHLSBELC 2165
Db 1741 LL-----EHRATYAKQOESKSHOSASLSKQSVASQEVQVOMLEDTYOVANVQMSMKDKLR 1796
Qy 2166 TYTTELQMLKQOKEEDINNKL--AEKYEVDELLOHLSLKEOLODQLOMELRNEKRLAYEL 2223
Db 1797 VTQOEKALQOEVMSLHKLOLQONAGKSMAPETIHPISGLHNO---QKRLSWPKLDH--- 1849
Qy 2224 CEKMDIMEKISIVYRLMQRNPOQOEDVARMILLESRNOEIOELMKISAVYSEQHTL 2293
Db 1850 -----LMNEEQ-----LMOENRLOQVMQVNTAEYLHSREKV 1883
Qy 2284 SSISSLEQKETEAKKH--CMINIKESLSTLSRFSGLQTEHYVLN---TOLQTLKLPK 2338
Db 1884 ROLESNLLPRKHQHNINSGIMNPLEQEKLSLKRECOQFOQEOGSAANKVQMSLSBELE 1943
Qy 2339 VYRTAAVKEHSLIKDYEKDLAAEQKRDHDLRLQLOCLEQHG-----RKWSDSASEEL 2392

```

```

Db 1944 TTH-----LENEGILKK-----KQVLEDOLEMOHLNSTATPPSPPHAW-DLOLLQO 1989
Qy 2393 KFCIEFLNELLFKKANIIOVSQODDFSEQVPLNQVSTLOEELHKKGMOMLEERGDL 2452
Db 1990 QACWVPREQLOLQORLOLQ-----ERIQHLOELELEN----- 2024
Qy 2453 HYDAKLSSEGWOENRRIASTIOLTLRLKRAVQOSKIOREITYVLANOFEAKLOEKKOEN 2512
Db 2025 -TSEFTNPQONQ-----QVLT-----VMEERIMEVEQKLVKRLDQEKVQOK 2068
Qy 2513 ELKRRMEHGBSASVMEENARLLGIKTYODESKLQSKIKMLENL-----NLVD 2565
Db 2069 FOICKNTKADAMVADLYVENAOLKALEVTEQROKTAERKNYLLKEKIASLSNIVRN 2125

RESULT 11
US-09-727-384-6
; Sequence 6, Application US/09727384
; Patent No. US20020098511A1
; GENERAL INFORMATION:
; APPLICANT: Myriad Genetics, Inc.
; APPLICANT: Helichman, Karen
; APPLICANT: Climbora, Daniel M.
; APPLICANT: Mauck, Kimberly
; APPLICANT: Bartel, Paul L.
; TITLE OF INVENTION: Protein-Protein Interactions
; FILE REFERENCE: 2318-271
; CURRENT APPLICATION NUMBER: US/09/727, 384
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 60/168,377
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: US 60/168,379
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: US 60/185,056
; PRIOR FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 2139
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-727-384-6

Query Match 5.4%; Score 794; DB 10; Length 2139;
Best Local Similarity 20.9%; Pred. No. 8.7e-23;
Matches 527; Conservative 444; Mismatches 912; Indels 634; Gaps 113;

Qy 271 KKLSDGQAGF---INRDSKLR-----ILQNSLGNKATVILCTTPVS 313
Db 21 KASDDVQVSGFHRKIQHVKNELCMLSLEAVAPVLOQTLLQDNLGR-----VH 69
Qy 314 PDELTLQPASTAKHVRNTPHNEVLDEALL-----KRY-RKEILDKKOLENL 363
Db 70 FDFKREAL-LLIISRLTSENHEHFE--PDCSLEQPKYVGGKYGRRSJPFEQSEVEEF 126
Qy 364 -----ESSSEKQAMAKKEH--TQLLAEIKOLHKEERIRIHLNINIVASSQES 411
Db 127 PEVTVIPLDEARPSIIPAGDSEHWKTO--RSEYEABGQLRFPVNPDLNANQSGSS 183
Qy 412 -QODQVRRKRRVYVWAPGKIQNSLHAGVSDFDMLSRIPNFSKKAFFSMPSPFETDSDS 470
Db 184 PQDQ-----WIEKIQEVEDCIGITRDGLNR-----KRLVSCIGQYGLQNVGDGE 228
Qy 471 VCTEF---SPFDALSMMD-----SNGIDAENKLAKVTHREKTSLSHQSHIDPGQ---- 517
Db 229 MLEEVFNLDPDGTMSVEDFEYGLFKNGKSL--TPASVTPYRLKRLHLSMQSFDESGR 285
Qy 518 -----ISDSVQFDSKQENQLOLYLPKDSGDMACRKASFEKETSLSQOQ-LQSKBEKK 570
Db 286 TTSSANTSTIGFR-----VFSCLDDG-----MGHASYERILIDQOEBEISQSLIK 333
Qy 571 ELVQSFELKIALEBEOLSVKAKNLEWNTNSREH-----SINAEVQDVER--EYVRKEMS 623

```

```

Db 334 ALDFSLDGNINITE--LTTALENELVTKNSHQALASAKAETIRHLEEDVDVANEKEK 391
Qy 624 VLGDSDYMASNSDLODSDVGKRLSSSHDECIEHRKMLEQIYDL--EFIEMLNKKSEN 661
Db 392 LRSDDLRAEKRLKSLMASEVD-----DHNAIERRENYNLRKLDGEYKKEKRIALAKHLELR 445
Qy 682 DKKSSEODPMESIOCEALMAEKANALEALMRONFONILLENELKREIALDERSLK 741
Db 446 ERQOIIIOQAKORLEBOL--EKATEEN--YIRDR-----LALSUK 484
Qy 742 ENOETNEFELKEKOTKEHBAOLIHIGSLUKLVENAMYNOMLEEDLETKRLKEOET 801
Db 485 ENSR-----LENE-----LENAE----- 498
Qy 802 QIALELKRADNLOKVRNFDLSVSGDSEKLCIEIFOLKOSLSDAVAVTRDAKESFLR 861
Db 499 KLAEYENLTKLORLNLEN--LAERFGDLDPSAEPFLQOERLTO--MRNEYRQCRVLQ 554
Qy 862 SENLELEKEMEDTSNMVNOEKAAFLFEKOLETEKSNYKMEAD--LOKELQSAFNEIN 918
Db 555 DOYDELOSELEF--YRAGRVL-----RLPLKNSPSEVEANSGIPEHGLSGEEN 605
Qy 919 YLNGLIAGV-----PRDILS--RVLEKRVSEFSQOLEKALE--KNALEN--EVTCL 966
Db 606 PLMMSIEALVIBOMKEOHHRDIOCLLELEDEKVRHYEKOLDETVVSCKKAQENMKQRHE 665
Qy 967 SEKKFIPNVECEKNOISKASEIIMLKOGESASISKOETIMOQSOQLDOLDEVT 1026
Db 666 NEKRILEKQISDLKNETALQOGAAYLK--EAHEAT--CRHE--BEKOLQVLEBEKT 719
Qy 1027 HQSKVOQTEEOYLEMKKMHDLFEKYIRNKSSEDLLREMNILKQTMESVE-----YK 1080
Db 720 HQEKRLQHE--MELKARLTOQAOSFEREREGLQSSAMTEKVRGULQOELQHEQOEL 777
Qy 1081 IADTKHELETITDKQOLHEKRYFQAMOTIPTIPPLSDSPSKLVBGNSSDPIEIN- 1139
Db 778 SLVEKHTLE-----KEEL--RKLELKHQ-----RELQEBREKMEKTECN 815
Qy 1140 -----DYHNLIATATERNNIMWC--LETENSLKEQYIDLNTQLOSLQASIKESDL 1189
Db 816 RNSQIENQOSDQKTER-----CESALOSLEGKRYQELKIDQO-----QREKSOQ 864
Qy 1190 QKPKODL--ESEGKLLLEMLIKGLHJDSQSIKQLQENLEVENTEKLQLOEMKNITI 1247
Db 865 EEFKEDLTQECABAQOELLKTKETKETSJVLQEREMLE-----KTYKHLNMMV 916
Qy 1248 ERRELQTNEDLKAHDSLKODLSENI EOSIETQDELRAA--QELRPQKOLVDSFQOLL 1306
Db 917 ERQOLLQDDEDLKNVSETOOSLSDOL-----ELKSSHRELEHREVL----- 961
Qy 1307 DCSVGJISSP-----NHDAVANOEVKVS--LGEVNSLOSIMLRGERDELQTSCK 1351
Db 962 -QOAGSBDQASQRLERLEMEHQ--ERQEMMSKMLAMENHKTCTADAREKEMST--- 1016
Qy 1352 ALVSELELLRAHYKSVGENLEITTKLNGLER--ELIGKSE--ESEVLKSML---ENLKED 1405
Db 1017 ----EISRLQSKIKEMQ--QATSPLSMLQSGCYVIGEEVYGDALSLILQGEOLLE 1068
Qy 1406 NMLK-----KQOAEYSSKEMQSLSEVFSQKLVDEIVLKAQKAAERLETKD----- 1457
Db 1069 NEDVLLISLOPAHEQAVKEN-----VKMATEISRLOQRLQLEGLVWSSCIDLBP 1117
Qy 1458 -BDYF-----ELVQANTNLVCEKLEPLQADHEED-----SIDRS---E 1494
Db 1118 ABEFFONTABQEQFLOQNRKQVEGYTRRHVJSLDEDEVRDLGSTGSSVQORQVYKE 1177
Qy 1495 EMEIKVLG--EKLENOYLLERLOEKELELKNLELLOKEMETSVILKDLQOKLEESLSE 1553
Db 1178 ESBASVEGSELENS-----EETRTSEMELKNOISQLOEOLMMLCACCDASSEKKODLLE 1233
Qy 1554 NIILKENI-----DTTLKH--HSDTOAQLOKTOOEQLAKNLAIAASDNCPTIOERE 1603

```

```

Db 1234 VSVLKKKLKMLERIPASPKYKLLYEDVSRENDCLQELRMETRYDALEN-----NKE 1288
Qy 1604 TSADCVHPLEEKILLTEELHOQTNBOEKLHEKNLEQAOV-----LKCEVH 1653
Db 1289 LTAE-----VRLQDELEKMEVEYETFLSLEKSYDEKLENEGLNVLRLQOKIKR 1340
Qy 1654 IMKSITES-----KSSLSIQHEKHDTQOQLALQOQVYVQEKKELOQTEHETLAEV 1707
Db 1341 LQESVQORDDCQMEASLENLEIE--PDGNILQINQOTIEECYPRRASHVHIEBECKQEN 1397
Qy 1708 DILKENIEGLNFKNEAQKTKTEQCLNBNKLEDSOHRLOCEI--EE--IMKSIDKE 1763
Db 1398 OYLEGNTOL-----LEKKAHEIAMLGTOIOTHOERPVNOQVILEENTILLGPODK- 1449
Qy 1764 SALETIKESQKVINLQOEMEMVLEMEELKNSQRFVIAERDLODDLRESVMSIEROD 1823
Db 1450 -----HFQOHTTAELELEK-----LOELTKLERVATILVQKD 1486
Qy 1824 DLKKAQALQOQKQVOELTQISVLOEKISL-----ENOMLYNVAATVKEETLSER 1874
Db 1487 VLSHGEKE--BELKAMHNDIQITCEMQOKVELLRYSEKLOQENSLIRNEIT---TLNEE 1542
Qy 1875 DDLNOKOHLFSEITFLSLKKEKFPALQAEKDKADAAKRTIDITEKISNIEBOLQOA 1934
Db 1543 DSISNLK-----LGLNQSQOEMQKTEYVQENNAVQKWEVNLKQIS--ELKIKNOQ 1594
Qy 1935 TNLKET--LYERESLQCKROLANTHELETLKSKDLALGKMEQOEDEAANKVIALTEKX 1993
Db 1595 LQENLELSOKNS--ONQEKLOELNORLMBLQOKKEGN----- 1633
Qy 1994 SLSLEQINENVTTLKGEKKEFTYLQRP--SKOQSSOMELRESLTKDLOLEBAEK 2050
Db 1634 SALEER-----EOEKFNLEKLEERKCYQSSVLSLEALESEVYIQTIVQO 1680
Qy 2051 E--ISEATNEIKNL--TAKISSLEELIQNASILNEAVSERENLHNSQOAYSELOJS 2105
Db 1681 ENHLDELEKMKQLOHRCPDLSDFQOKISSVSYNEKLLKEKALESEBLSNCVDKLAKSS 1740
Qy 2106 LTLKSRDHAFQOSKRREKDEAVNKRASIAEIKITIKEMDEFBRSKSLQEOSSHLSSELQ 2165
Db 1741 LL--EHNLATMKQKQKMEHQSAKSLQSLVQASQKQVONLEDTQVGNVLQMSRMSDLR 1796
Qy 2166 TYKTELQMTKOQKEDINNLK--AEKYKAYDELLQHLUSLKQLODQOTMELRNEKLNTEL 2223
Db 1797 VYQOEKALQOEYMSLKHQLOMNGKSMAPETATHPSGLHNQ--QKRLSDMKLDH--- 1849
Qy 2224 CERKMDIMEKISVLRIMQNEPOOEDDVAERMDILFSRMOEIOLEKISAVYSEQHTLL 2283
Db 1850 -----LMBEBO-----LMOENERLQTMQNTKALETHREKRY 1883
Qy 2284 SLSSELQKETEBAKH--CMLNIKESLSSLSKSPGSLQTEHYKLN--TOLQTLNKKR 2338
Db 1884 QOESNLPLKQHOKHLNPSGMNPTBQEKLSLRECDQFQKQSPARKVYSQNMSSLOELE 1943
Qy 2339 VYKRTAAVEDSHLIDYKDLAABQKRHDELRLQLOCLEQHS-----RKSDSASEEL 2392
Db 1944 TJH-----LENBGLKK-----KQVKLDEQIMQOHRSTATPSPSHAW--DIQLLQ 1989
Qy 2393 KFCSEIFMLLEFKRANITIOSVDDPSEVOYVFNQVSGTLOEBLEKQKFMQMLEFGDL 2452
Db 1990 QACPMVPRQOTLOLQOLQOA-----ERINQHLQEBELNR----- 2024
Qy 2453 HDQAKLSEGMOENRRIASTIQLTLRLKAVQOSKIQOREITVYVLQWFAKLOEKKEQNK 2512
Db 2025 -TSEMTPOGNO-----QLYV-----VMEERMLEVQOKLKVYRLLQEKYNQK 2068
Qy 2513 ELMRMEHNGPSASVMEENARLLGLTKVODESKRLQSRIMKLENEI---NLVQD 2565
Db 2069 EOLCKNTKADAMVQDLYVENAQILKALEVTQEBORQTAERKNVLEBKIASLSNIVRN 2125

```

RESULT 12
US-09-820-843A-73


```

; Sequence 73, Application US/09820843A
; Publication No. US2003039963A1
; GENERAL INFORMATION:
; APPLICANT: Council of Scientific and Industrial Research
; TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE PROTEI
; TITLE OF INVENTION: USEFUL AS ANTI-INFECTIVES
; FILE REFERENCE: 063915
; CURRENT APPLICATION NUMBER: US/09/820,843A
; CURRENT FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: Patent version 3.0
; SEQ ID NO 73
; LENGTH: 1805
; TYPE: PR
; ORGANISM: M. genitalium
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: g11045905
US-09-820-843A-73

```

Query Match 5.2%; Score 761; DB 9; Length 1805;

Best Local Similarity 21.2%; Pred. No. 1.3e-21; Matches 437; Conservative 392; Mismatches 768; Indels 465; Gaps 86;

```

QY 678 KSENDKOKS-SEODFMESIQLCEAIMEKANALEELALMRDNFDNI-----ILENETLK 730
DB 24 QSEVEKRSKTDVVKIEN-----QLKEIKSLSEDELNKLGKAKQAEDNPELD 71
QY 731 READERSLKENQETNEFEILEKETOKHEBAQLIHEIGSLIKLV-----ENAEWY 781
DB 72 KKHNLVEVDL--NRLVNEYK--NFQOKNHVWDVSELDNTFRYKNELTRIQOENADEFL 127
QY 782 NQ-----NEEDLEFTKILKEOELQIOLAEIRKADYLQKVRN-----EDLSVSM 826
DB 128 NSKANTANFOANTHNKLNDFHLEINOMQNTINLNQKINQNMILIDNVALLONPNTV 187
QY 827 GDSEKCEELFQKQSLSDAEAVTRDAQKESFLRENTL--ELKEKEDYSNMYNOKEKA 884
DB 188 EKNYLLNVDIOLYNELDQLE-----NQKRLSTIEYENTYELVSADENLQYVENIDQN 242
QY 885 ASLPEKOLEEKSNTKMEADLOKELQSAFNEIN-----YLNGLLAGVPRDL 933
DB 243 QIOFKHQYQYRDLSELERKIQTLQKQELVDKESALRYKIDDAFYINARLA----- 294
QY 934 SRVLEKKEVSEFKOL--EKALKEKNA--LENEYTCLSEYK--FLPNEVECKLNQISKASE 988
DB 295 -----ELDDVAKQLSFQDQITKONAHVEDKLVALNKEKORLNTQAEAFNLRQSAALI 347
QY 989 EIMLLKQEGHSASIIIS--KQEIIMQEQSEQIQLTDEVYTHQSKVQOOTEQYL---EMK 1043
DB 348 DINKLOQENELFAKHLEHQNEFQKQSDSLKLETEYKALQHKINERKNESATKSEELL 407
QY 1044 KMHDDLEKTIKNSSEADLLIREMENLKGITWESVEKIAQTKHLEETIRKQELIHEKK 1103
DB 408 NOERLEPER-----RREDITLL-----TQASLEY--EHQRESSQLKDKQ---NEVK 449
QY 1104 YFQAMQITPIPIPLSDSLPSKLVESQSDPIEINDYHNLIALATERNNIMVLETERN 1163
DB 450 QHQO-----NLEYAKKE-----LDKERN 467
QY 1164 SLAEQ-VIDINTOLQSLQASIEKSDLO-----KPKQDLEGEVYKLLLEMLLKGHLTD 1216
DB 468 LLDQOKVYDSEBAIFQLEKVAQERKELELYLVKKQKQDQENE--LLFFEKQLQKQAD 525
QY 1217 SOLSIEKQLENELEVTKLQ-----TLODEKMNITIERNELOTNEDLK-----AEHDSL 1266
DB 526 FENLELEKQOELFEAKHALERSFTIKLEDKEDLTKAQOIANEFSQKTDKSKSADELM 585
QY 1267 KQDLSNEIOSIE-----TODELRAOQELREOK-----QLVDSRQOLD 1307
DB 586 LQNTYEMLOQEKQKLFQERTYFFERNNAVLNRLQOKREELLQOKETTDOLTKSSEQEL- 644
QY 1308 CSVGSISSPNH--DAVANOEKVSLGEVNSLOSEML--RGERDELQTSCKAL--VSELELLRAH 1363

```

```

DB 645 -----INQREHKEVLASVEK-----QKELGKQLQDQFSQTSLNASKNLAEERMA--- 688
QY 1364 VKSVGEMLLETTKUNGLEKELKSESEVYKSLMLENLAKEDNNKLYKQAEAYSKNEQF 1433
DB 689 IKFKE-----KELEATEKQILDVNNVAEYIOADLQJLQNSLOERSELQANQORLADF 741
QY 1424 -----SLEEVFSGSKQVYDELEVYKLAQKAAEERLEIKDQYFELVQNTANTNVEKLE 1477
DB 742 HNDLSKLTNEYELSLQKLOLQETLEAN-----QKQHSIQONAYF-----EGELD 786
QY 1478 TPLQADHEEDSIDRSEMEIKVLEKLE-ERNQYLLERLOEKL-----ELSNKL---EI 1538
DB 787 -KLNRKQAFNLNRKKQYMEVDALIKQRLSDRHQAL--NMQQAELDRKTHELNNAFLNMDA 843
QY 1529 LQKEMETSVLLKDDLOKLE-----SLSENIILKENIDTTLKHSIDTQAOLOK----- 1577
DB 844 DQKSLQDOLATVKEQOKLIDERSALLKQREFAENVAGFRHMSNKTSQLQKIVELYLRK 903
QY 1578 -TQOELQAKMLATASDNCPTIOEKETSADCVHPLEKILLTLEELHQ---KTNEOKL 1633
DB 904 QSEQOTKETELKTAFSQLOQDYOFELQKD-----QEFQITAKQRELDKL 950
QY 1634 LHERNELQAOVELKCEYEHLMKSMIESKSSLESLOHEKHDTQOLLALQKQOMOVYTOEK 1693
DB 951 AEKNQYK-----LELDNRFQALQNGKQOTYQAOLELEREHQHLMQ 993
QY 1694 KELOQTHEHTLAEDVHLKENIELGLNFKNEAQKTKKQCLNENKEL---EQSOHRLQ 1750
DB 994 TAFQANESILKQREOLKQKQIA--FYHELK--RNQFALALGKRLFAKEAQDOORQD 1047
QY 1751 EIEELMKSLDKESALETKSEBQKVINLQNMEMVMLEMELKNSORTVYAEBOQDQD 1810
DB 1048 EIMNRFQOFEKTYDFDEDAKR-----ELELEKIRSL---SOSAVE 1087
QY 1811 LRESVMSIETQDDLRKAQALQOQKQVQELTQSIQVLAQKISLENNQMLYNATVYET 1870
DB 1088 LERRERKATDFMLNKVQHTQINRD--QNSQI-----ROFLERKNFORS----- 1134
QY 1871 LSRDQDLNOSQHLFSELET--SLSLKEKEFALQAKQKADAAKTIIDTEKISNIEE 1928
DB 1135 ---NEANAKRAFILKRLRSFASNLKQEKALATQKLEFDRDQK----- 1177
QY 1929 QLLQOAT-NLKETLYEREST--IQCKEOL--ALNTEHLRETLKSKDALGKMOERDEAANK 1985
DB 1178 KELQOATIQLOQFKPEKONFPIEKQROLVAIKQ--CEKLEDEKALNQLKVELKNSQT 1235
QY 1986 VIALTEKMSLEEOINENVTTLKEGEGEKEFYQLQPSKQOS--SQMELESISLTKYDQ 2044
DB 1236 YLANKNKAEYQOQLOQKQYTMILDKLEMLERTQDQDKHRSIFARLTQFANDLRFERKQ 1295
QY 2045 LBAKELISEANTENIKNTLTAISSLEBEILLQNASILNVAVERENLRISKQYL----- 2058
DB 1296 LTKAQYVDDDNRLKLENERNHLFSLNETERRRAVLQEOISYFQKQOAVDALASHKE 1355
QY 2099 -----SELEQSLTKRSR---DHAFOSKREKDEAVAKISLAEIKILTKENDEPDS 2149
DB 1356 VKKKEGELQKLVLELERTKTLNNDPAKFSQREFEQORLKLLELOKTLOTQNSNNFK 1415
QY 2150 KESIQEQSSHLSEELCYKTELQMLKQOKEDINNKLAKEYEVDLLOHSLSEKQLOQI 2209
DB 1416 TKAIOETIEN-----SYKRGMEELNFOKKRFD-----KNKSRLVEYFPRKMDETERK 1461
QY 2210 QMELNRKLTNYELCEKMDINERKISVLRMQN-----EPQOEDDVARMIDLESRN 2262
DB 1462 ESQV---KLVLKETQORANLLEQAANKLNTENKNTIDFKERKLKAFKQVDDQDISTNKR 1518
QY 2263 QEIOELMEK---ISAVYSEQHTLLSSLELOKTEAHKHCMILNKE---SLSSYTSRS 2315
DB 1519 KELNELLENKLLQOOSLIERBRAINSKDSLANKIETITKR-QLNDKEMRYRLVDRKRLA 1577
QY 2316 FGLQTEHVKLNTQL-----QTLNKKFVYRYTAAYVEDSHLINDYKDLAAEQKRDEL 2370

```

Db 1578 EOKYQTEINRLTQTEPSEKODIKNEPPLFKI-----NGNDMAFPYLWLPYPOKODDN 1633
 QY 2371 RLQLOCEHGRKMSDASSELKCELEEFINEL--LEKRNIIQSODDESEVQFVLNOV 2428
 Db 1634 TLQLOLFE-----OQLOPMOORVENEINELRNHLEKILKD--QIQ----- 1674
 QY 2429 GSTLOEELHKKGFOMWLEFEFGDLHYDAKKLSGMOQENRIASTIQLLRKA--VWQ 2486
 Db 1675 ---LESQLNKQ-----SEFSKVESMEKLEKTESRLMDPDKINYLTKVNOHNTQY 1725
 QY 2487 SKIQREITVYLN---QFEAKLOEKKQNKELMRHEHGPSASVMEENARLLGLIKTV 2542
 Db 1726 PSSYQPPSPYQSDSKQOLFRIOLEKRON--LEQOQFOPAP--AVVOQPSFAAPNITK-- 1780
 QY 2543 ODESKKLOSRIKMLENELLVK 2564
 Db 1781 -----QOOIQADLNMEINNIK 1795

RESULT 13
 US-09-885-535-4
 ; Sequence 4, Application US/09885535
 ; Patent No. US20020104105A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Myriad Genetics, Inc.
 ; APPLICANT: Helichman, Karen
 ; APPLICANT: Bartel, Paul L.
 ; TITLE OF INVENTION: Protein-Protein Interactions
 ; FILE REFERENCE: 2318-266-II
 ; CURRENT APPLICATION NUMBER: US/09/885,535
 ; CURRENT FILING DATE: 2001-06-21
 ; PRIOR APPLICATION NUMBER: US 60/213, 245
 ; PRIOR FILING DATE: 2000-06-22
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 4
 ; LENGTH: 2835
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-885-535-4

Query Match 4.9%; Score 720.5; DB 10; Length 2835;
 Best Local Similarity 18.8%; Pred. No. 7.3e-20;
 Matches 515; Conservative 508; Mismatches 954; Indels 761; Gaps 111;

QY 315 DETLSTLOPASTAKHVRNTPHNVENVLDDEALLKRYKEIILDKKOLENSESSETKAQM 374
 Db 117 DNKLSSTLAVST-----HPDAMNQOLETAQKMKQELIQOEKKQIKVAQALCE-DLSAL 167
 QY 375 AKREHQ-----LAEIKQLHEREDRIWHLTNIVASSQESQDOQRYKRRRTW 425
 Db 168 VKEEYLKAELSRQELILKSFQDVEQKAEHNVHQLQS-ACASSHPQMSRDFQ----AM 222
 QY 426 AFGK--IQNSLHAGVSDPDLRLPGNFSKAKAFSDMPSPEID--DSVGTSEFSQFDA 481
 Db 223 LQTKREQKSH-----PISAKDLVLESILIDHDKDFST 256
 QY 482 L---SMADNGSIDAEMNLASKYTHREKTSLHOSMIDFGQISDVQFHDSKRENOQLYLPK 538
 Db 257 LTAQSHMYEKTIAEGENLLKTKQSEKALQLOLNTIKTNWPTFNQVYKEREKTL----- 311
 QY 539 DSGDMAECKKASPEKITLQO-----QLOSKKEEKELVQSFE-----LKAIELE 584
 Db 312 -----KESLEKALKYKEQVETLMPIDKQNNLEIKFCIDPAEENSIAKLKSIQ 362
 QY 585 EQLSYAKNLLEWYNSREH-----SINAEVOTDVEKEVYRKESVYLGDSGYNAASNDLQD 639
 Db 363 KEMDQHFQGNVELLNTNANSLVCELDKRYVYDENKSLQK-----VDWYT 408
 QY 640 SSVDGKRSLSSHDECIEHRRMLQKIVDLEEFIEINLKKSENDKQSSQO--DFMESIQLC 698
 Db 409 EQLHSKKF-----CLEN---WTQKFKERQF-----VSKESKRLQCAKAKQQLDIDHSL--- 452

QY 699 EAIMAKKANALEELALMRQNFNIILENETIKREIADLERLSUKENQETNEFLEKETQK 758
 Db 453 -GQATSNKTYITMLQTOQSLQALKHQVDAKLADLVASDSGTSQV--LLOVETIA 510
 QY 759 EHEAQLIHEIGSLKLVEN-----AEYQNQLEEDLETKTLKRLQEIQLAEL 806
 Db 511 QEHSTLSQVDEKCSFLFTRKIQIGHFQNTIREMFQAFEPDELDSS-----MAYV 561
 QY 807 KRRADLQK---VRNF-----DLSVSGDSEKLCEITQLOKQISLDAVAVTRDQ--K 855
 Db 562 GRDAETLOKQETIKAFLEKLEALMASNDNANKTCMMALTEBTPDLVIGIRDEALSK 621
 QY 856 EESFLESEMLEKEMEDT-----SNWYNOKERAASTFEKO-----LETEKSN-- 898
 Db 622 QCNKLLDRQARQREYEGITKRLFEETYSKLESTILQAAEHESQSGVGHETETINQ 681
 QY 899 ---YKMEADLOKELOSAFENEIYL--NGLLAGKVPRLDLSRYE-----LEKKV 942
 Db 682 LNMFKYFOKEEIEPILOQKQDVNMLGQGLQSAKSTSTQGLEHDLDDVANARKTLLNKV 741
 QY 943 SEFSKOLEKAL-----EKNALENEVTCSEYKFLPNEVECLNQISKASEEIMLKQEGE 998
 Db 742 AQRRAQLQELALHCGRFQDALE-----SLSMWDTIELVANQ--KPSAEKRYK--- 790
 QY 999 HSASIIKQELIMQESQELQLTDEVTHTQSKVOQTEQOYLEMKMHDLFEKYIRNKS 1058
 Db 791 -----AQIQQKL-LQRLDD-----RRS 808
 QY 1059 EAEDLIREENLKGWSEV-YKI-----ADTHLELEETTRDKQOLH 1100
 Db 809 TWEVIRREGEKIATTAEPDKVKILQSLDSRWEALINKAETRRQLEGISVVAQOPH 868
 QY 1101 EKKYFQAMQITFPIPLSDSLPPSKLYGNSQDPIEINDYNLALATERNNIWC--L 1158
 Db 869 E-----TLEP-----LNEW-----LTIETKRLVNCERI 891
 QY 1159 EFERNSLKEQVIDLNTQLOSLQASQIENSQDLQPRQDLEEGEVKLLLEMLKGLHTDSQ 1218
 Db 892 GTQASKLEEQIAQ-----HKVLQEDILRRQNVQQA---LNLGIELIKQTTGDEV 938
 QY 1219 LSIENKQLENLEVTREKQLQLOEEMKNITIERNEIQTNEDLKAHNSDKODISENBSI 1278
 Db 939 LTIQ-----KLEIKARYDIT-----KISTDAKTLQAL 970
 QY 1279 ETQDELRAAQEELRQKQVLD-----SFRQOLT---DCSVGISPPNH--DAVANOEVSGE 1330
 Db 971 QIARLRHSHHELCWMLQVEVELLSYEQVLKGEASQAGQRRPELKEAKNNALLDS 1030
 QY 1331 VNSLOSMLRGERDELQTSKALVSELELLRAH--VKSVEGENLE-----ITKRLNG 1380
 Db 1031 LNEVSSALL-----ELVPMWRAREGLEKKAEDNDRYRLVSPDTIQKVEE 1074
 QY 1381 LEKEILGSESEEVLSKMLENKEDNNKL-----KEQAEYSKSENGFSLE----- 1426
 Db 1075 IDAALIRSQQFQDAADAELSWITTEKRLMSGIDIRLBDQTSQAOLQVQKTFTEILRHK 1134
 QY 1427 ---EVSQSOKLV-----DEIEVLKAQ---LKAARELEIKDRDYELVQNTANLVE 1473
 Db 1135 DIIDDLVNSGKHIMACSEBEQSKMKKLDKVLKYNVDITQINSRYQLDBEA--QSLVN 1192
 QY 1474 GKLEF-----PIADHEDSDIDRSEMEIKYLGKLERNQYLLERLOEER----- 1519
 Db 1193 QWETYEELMPWLTETQSIISQLPAPALEYETLRQOOQEHQRLRELIAEHPHIDKMNKT 1252
 QY 1520 ---LELS--NKLEITLOKEMETSVL---LKQDLOK---LELSISENIIKENID----- 1562
 Db 1253 GPQILELSPGBSFSIQEKYVAADTLYSOIKEDVKKRAVALDAIASQSQFQHDKIDQILLES 1312
 QY 1563 ---TLKHSDDTQAOLOKTOQLOLAKNLALIASDNCPTIQ-----E 1601
 Db 1313 LERIVERLRQPPSISAEVETIKQEQISENNKNSVDEMEKIQPIYETPLKQNGEIMARSGTD 1372
 QY 1602 KETSADCVHPILEKILILLTEBELHQKTNEQEKILHEKNLEBOAOVELKCEVHLMKSMIES 1661


```

1373 KDLSAKVQKLDQVAFIWEIHIHLVEERAKLLDVMEAE---KFCMDHSLVITIKDT 1429
1662 KSLIESIQHEKHDTQOQLLAKQOMVYVTOEKKELOQTHEHLVIAVDHLKENIELGLNFK 1721
1430 QDFIRLEDEGIDPS-----VVKOO-----QEAETIRREIDGLQELDIVINLG 1474
1722 NEMQKTKTKQCLLNKEKLEQSOHRLOCEIEELMKSJKOKESALFELTKSEQVIVINQ 1781
1475 SE-----LIAAGEPDKPIYKKSIDELNSAMDSLNKAMKDR---IDKLEEMQAAVOYOD 1526
1782 EMEVYMLEMEELKNSQRTVIAERDLODDRESYEMSIETODLRKQOEALQOQKQOE 1841
1527 GIGAV-EDWDIANGK---LASMSPIGTDL-EIVKQOE---ELKQFKESEYQOQIEMER 1578
1842 LFSQISVLOERIS-----LLENQMLYNAVYKETSERDILNOSK-----1881
1579 LNHQAEILLKRYTEESDKHTVODPLMELKILW-----DSLERRIINROHLEGALLALG 1632
1882 --OHLFSEI-----ETLSLSLKEK-----EFALBEAKDKADAARKTIDITEKI 1923
1633 QFQHALDELALWLTHTGLLSEOKPVGGDKPAITEELAKHHVLONDVLAHOSTVEAVNKA 1692
1924 SN--IEBQLQOATNLKETLYERESLIQCKEQOALNTEHLETL-----KSKD 1969
1693 GNDLIESSAGEEASNLONKL---EVLNQRQNVLEKTEQRKQOQLDGLARQKGFHEIED 1749
1970 LAIGKQOEDEBAANKYI-ALTEKMSSLEQINENVTTLKEGESEKETFYIQRSKOOS 2028
1750 LOOMLDTHERHLLASKPLGLPE---TAKBOLNVMEECAFAKEKEITY-----KSLMQKG 1802
2029 SOMELERESLKTQOLOEAEKEISEADNEIKNLTAKISSLEETILONASTILNVAVSRE 2088
1803 QOM-LARCKRSATNTIDQ-----DINNLEKEMESVETKLNKTKTLEALINAM 1850
2089 NLRHSKOOLVSELEQSLITLK--SRDAFAQSKREKDEAVNKIASLAEIKILTKREDEF 2146
1851 EFHNSLODFINMLTQABQTLNVAASPSLI-----LDTVLFQIDEKHVFANEVNSH 1900
2147 RDSKESIQEBSHL-----SEELCTYTELOMLKQOKEDINNAEKVKEVDLOHL-- 2199
1901 REQIETLDKGTHTLYESOKODVVLINLNLISVQSRWEKVYQRLVERGRSILDAARKRAQ 1960
2200 -----SSLKQOLOQLOELNREKLRNYELCEKMOIMEKEISVLRLQNEPOQOEDVAER 2254
1961 FHEKMSLMELESESLSE---LEIANDPKIKTOLA-----QKFEPOKSLGAK 2009
2255 MDLIESHNOETLOELMERTISAVYSEQHTLSLSELOKETEAHKKMLNTIKESLSLTSR 2314
2010 HSYVDJTNRTGRSLKEKTS--LADDNLKLDMDLSELDKMOT--ICGKSVEROKMLEAL 2065
2315 SFGSLQTEHAKLNTQLOTLNKFVYVRYTAAVKEDHSLINDYENDLAEQKRHDELROL 2374
2066 LFSQGFDD-----ALQALD--WLYRV-----EPQLEDOQVPHGDIDILVM 2103
2375 QCLEOHGRKMSDSASEELKCEIEFLNELFKKANIIOQVQDDSEVOVFLNQVSTLOE 2434
2104 NLIDNH-----KAFQK 2114
2435 ELEHKKCFQWLEEFGLHVDAKKLSFGMOENRRIASTIOLTKRLKAVVOSKIQRETT 2494
2115 ELGKRTSSVQALKR-----SARELLIGSRDSSGVMYQWQELSTRMETVCALISKO-- 2166
2495 VYLNQFAKIOEKKEONK-----ELMRBHEHGPSASVAMEENARLIGLITVQOD 2544
2167 ---TRLEPAALRQAEFISVYVHALLEWLAEBQTLRFGH---VLPDDDEA---LRTLID 2215
2545 ESKLQSRIKLENELNLVDDAMHKEKVALIQQDKLSRNASAEALNMAQYKLTQKODNL 2604
2216 QHKFEMKLEKRAELN-----KATPMGDYVALAICHPSITTIKIMIIILIRARF 2264
2605 QAAKKEIENLOKVAK---GAVPYKEETDNL-----KTKVVKIEMEKIKYKSKATDOI 2654

```

```

Db 2265 EEVLAMAKHQOURLASALAGLIAKQELLEALLAWLQNAETTLTDKREVI-----POEI 2318
Qy 2655 AYLSCEDEKREGLRLKEELRRQADNDITVCYKPKDQKASTFPYVCGGSGGIYSTAM 2714
Db 2319 EEWALIAEHQ-----TEMEETRKQPDVK---VTKYTKRRAAPSS-----2358
Qy 2715 LVLOSEKALERELSHYKKRYHNLSTRWSSSEDRKTK 2752
Db 2359 --LOSHIPVLDKGRAG--RKRPASSLYPSSGQTOIETR 2393

RESULT 14
US-09-927-597-4
; Sequence 4, Application US/09927597
; Publication NO. US20030032018A1
; GENERAL INFORMATION:
; APPLICANT: Malik, Fady
; APPLICANT: Berard, Christophe
; APPLICANT: Freedman, Richard
; APPLICANT: Craven, Andrew
; APPLICANT: Sakowicz, Roman
; APPLICANT: Hartman, James
; TITLE OF INVENTION: Human smooth muscle myosin heavy chain
; FILE REFERENCE: CYTOPO18
; CURRENT APPLICATION NUMBER: US/09/927,597
; CURRENT FILING DATE: 2001-08-10
; NUMBER OF SEQ. ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1979
; TYPE: PRT
; ORGANISM: Human
US-09-927-597-4

Query Match
; 4.98; Score 719.5; DB 9; Length 1979;
Best Local Similarity 19.88; Pred. No. 5.5e-20;
Matches 462; Conservative 428; Mismatches 847; Indels 591; Gaps 93;

Qy 17 IOROGDQANLQW-----KAGNNTISOVDGTSFNFDRVFNHSHESQIYOELAVPII 69
Db 50 IKKEGDEVVVELVENGKKTATVGKDDIOKNPK-----FSKYEDMAELCLNEASVL 102
Qy 70 RSALQGY-NGTIFAYGQTSQKTYTMGTSPNSLGIIPQAOIEVF--KIIQEIPIRNEFLR 126
Db 103 HNLKERTVSGILTYI---SGLFCVYVNPYKHLPIYSKIDYMKKGRHMP-----151
Qy 127 VSYMETIYETVKLLCDDRKKPLPIREDFENRNVY---ADLTEELVMPEN--VYQW 179
Db 152 --PHIY-----AIADAVRSMLODRD--QSILCTGESGAGKTEENTKKYIOYLAVVAS 200
Qy 180 IKKGKKNRH-----YGE-----TKMNBHSSRSHTIPRMIVESR 212
Db 201 SHKGGKDDISTQGPSFAYGELKOLLOANPLLEAFNAKTYKANDSSFGKFIH-----255
Qy 213 DRNDPTNSEND--GAVVWSHLNLVDLAGSERASQGAEGVRLKEGCNINSLPLGVI 270
Db 256 -----NFDVGYIYGANIEIYILEKSRAIRQARDE-----RPHFIETYMI 295
Qy 271 -----KRLSDQAGGFIVY-----RQSKLTRIIONSIGNAKTVIICTIPV 312
Db 296 AGAREKMSDLDLLEGFNNYTLNSNGEVPPIPAADDEMO-----ETVEAMALINGF 345
Qy 313 SFDETLSTLOFASTA-----KHVRNTPRVNNEVLDEALLK-----RYRKEIL 354
Db 346 SEEDOLSLTKYVSSYLQNLGNTVFKKERNTDAS--MPDNTAOKYCHLMGINVTDFTRSL 404
Qy 355 DLK-KOLENLESSSETKAQ-----AMAKEBHQLAEIKQLHKEREDRIWHLNTIVAS 407
Db 405 TPRIKVGDDVYQAKQTEQADFAVEALAKATYERLFR-----WILTRV-----447
Qy 408 SQESQODQVRKRRRYTAPKQIONSILASGVSDPDMLSRLPGNFSKAKKSDMPSP-- 465
Db 448 -----NKALDKTHROGASPLGIL-----DIAGEIIF 473

```

QY 466 EID--DSVCTEFS--FDDALSMDS-----NGIDAEMNLASKVTHREKTSILHQM 512
 DB 474 EVMSFELCINTYNEKIOQJFNHTMFLBEOBVEYOREGI--EWN-----F 515
 QY 513 IDFG-QIDSVDQF--HDSKRENOLOVLPKDSGMAEBCRKASFEKEITSLQOOLQSKEEK 569
 DB 516 IDGLDLOPCIELERNNPGLALLDECEWPKATDKSFVEKICLT--EQSGHPKQPK 573
 QY 570 KELVOSPELTAIELEBOLSKA-----KLEMTVNSRHSINA--EVOVDVEKVRKE 621
 DB 574 KOLDKTEFSITIHAGVNDYNASAMLTKNMDPLNDNTSLINASSDFVADLMDKVDNR-- 631
 QY 622 MSYLG--DSGYNASNSDLODSVDGKRLSSHDECIEHRKMLEOKIVLEEFIEITLAK--KS 679
 DB 632 --IVGLDQMAKMTSSLPASKTKKGFRTVGOL--YKEDLGKIMTLTIRMTTPFVACIT 687
 QY 680 ENDKQSSSEOD--FMESTQL--CEAIMEKANALEBELALMDNPNILLENETLKR--ETA-- 734
 DB 688 PNHKRSKGLDAFLVLEQLKNGV-----LEGIRICRGFPRIYVQFQFRQYETILAA 740
 QY 735 --DLERSLKEQNETNEFI--LEKETQKHEHQLHEISLKLKVENAM-----Y 781
 DB 741 NALPKGFMDKQACITMIRKALELDPNLYRGOSKIFERTGVLAHLEREDLKITDVIMAF 800
 QY 782 NQMLEEDLEKTKLKEOEIOLAEIRKADNLQKVRNPD-----LSYMGDSE 830
 DB 801 QAMCROGLAKRAKARQOOLAMKVIORNCAYIKLBNQWMLFTKVRLLQVOTRREE 860
 QY 831 KLC--BEIFOLKOSLDAEAVTDAQRECSFLSENLEKME--DTSNNYNOKERAST 887
 DB 861 MQAKRELOQTKERQOAEHELKELEQKHSQLETKENLLOEOLOAETELAEAEEMVRL 920
 QY 888 FEKOLETEKSNYKMEADLOKELQSANENIYNGILAGVPPDLRSVLE--KVYSEFS 946
 DB 921 AAKQOELELILH--EMERKEE-----EDRGQOLOAERKKAQOM 939
 QY 947 KOLEKALEENKALE-----NEVTCLESEKPLPNEVECLKNOISKASEMILMKOEGHSA 1001
 DB 960 LDLEOELEEEBAAROKLOLEKVTAAKIKLEDEILVMDQNNKLSERKLLEER----- 1014
 QY 1002 SITSKOEILMOESEOLOLTDDEVTHQOSVQOOTEBOYLEMKAMHDLFEKTRINSEAT 1061
 DB 1015 --ISDITTNAAEEERKANNTLKNKHEMSISEL--VRLKREE-----KSRBELE 1061
 QY 1062 DLTREMENTKGTESYEVKTAOTKHELEETIRKOEHLHKKKFQOMQOTIPPTPSDS 1121
 DB 1062 KLRKLE--GDASDFHEQIAD--LOAOIAELKMOJAKKEBLOA-----ALARLDE 1109
 QY 1122 LPPSKLVEGNSODPIEINDYHNLALATERNNIMVCLETERNSLKEQVIDLNTQLOSLQA 1181
 DB 1110 I-----AKNNML-----KIRELEGHISDLOEDLDSERA 1139
 QY 1182 QSTESDLOPKODLEGEVKKLLEMLKGLHTDSOLSTIEKLOLENTLEVEKLOTLQOE 1241
 DB 1140 --ARKNAEKQKRDGLG-----ELEALTELEEDT-----LDSTATQOELRAKRE 1181
 QY 1242 MKNITERNELQTNFEDLKAEHDSLKODLSENIEOSTIETODELRAQOELREKOLVDSF 1301
 DB 1182 --EVTYLUKRALDEETSHENOVOEMROKNAQAVEBELTEQOFRKAKANDKKNOLIEKE 1239
 QY 1302 ROOLDCSVGISPNHDVANOKESVLGEVNSLOSEMLRGE--RDELOTSCALVSELEL 1359
 DB 1240 NMDLAG--ELRYVLQAGQOEVHKKKKLEAOVOELOQOSKOSDERARAEINDVHKLQNEVES 1298
 QY 1360 LRAHVASVGENLEITKKLNGLEKEKELIGSESEVLSKSMENIKEDNNKIKEOAEYSK 1419
 DB 1299 VTGMLNEAREBEKAIKLKDVASLSOSL--ODTOELLQOEETROKLNVSTKIROLEERNSL 1355
 QY 1420 ENQFSLEEVYSSGOKLVEDEVVLKAOJKAEELEIKDRYFELVYQATANNVLVEGKLETP 1479
 DB 1356 QDO--LDEEMKAKONLEHRISTINLOJSDSKKLQ--DFASTVEA-----LEEGR-- 1401

QY 1480 LQADHEEDSIDRSSEMEIKVLEGEKLENOYLLERLOEKELELSNKLLELOKEMETSVLL 1539
 DB 1402 -----KFOKEIENINLOQYEKAAVADKLEKTNRNLQOELDDLVLDNRQL 1449
 QY 1540 KDDL--OQKLESLSLNTILKENIDITTLKHSDYQAOLOKTOEOLOAKNLAIAASDNC 1596
 DB 1450 VSNLEKROKRIQDOLLAE--EKNISRYADEAD--RAAE 1484
 QY 1597 PIQOEKTSADCVHPLEEKIILLTEELHOKTNOEKLHKNLEBOAQVLEKCEVEHIMK 1656
 DB 1485 --AREKET-----KALSTARAL-----EELKEKELEERKNKLMKEMEDVYS 1525
 QY 1657 SMIESKSSLELOHEKHDPEOOLALQOMQOYVTOEKKELQOTHEHILTAEVHILKENIEL 1716
 DB 1526 SKDQVGNVHLEKSKRALETOGHEMKTOELEEDELQATEDAKRLLEVWQMLKQOFER 1585
 QY 1717 GLNFKNQOQTKTEOCLNENKELEOSORLOCEIPELMK-----SLKDK 1762
 DB 1586 DLQARQONEE-----KROLOROLHEYTELEDEKROBALAAAKKKEGDKDL 1636
 QY 1763 E--SALETKESQKVINLQEMEMVMELEIKNSORTVTAERDOLQDDLRSEVMS 1818
 DB 1637 ELQADSAIKGEREAIKOLKRLQOMKDFORELEDARASROEITATAKENKAKSLADL 1696
 QY 1819 IETQDDLKKAQEAALQOQKQVQELTQISVLOEKISLLENQMLYNATVKEITLSEBDLN 1878
 DB 1697 MQOEDLAABERARKQADLEKELEAELA-----SLSGRNALQ 1735
 QY 1879 QSKOHLFSELETSLSKKEFPALQAEKADAKRTDITRKISN----- 1925
 DB 1736 DEKRRLREARLQOELEEEQGNMEA--MSDRVRAQOAOQLNSELATERSTAOKNE 1791
 QY 1926 --IEOOLQOATNLKETLYEESLQOKEQLALNTEHLRETKDQALGKMEQORDEAN 1984
 DB 1792 SARQOLERQONKEKRLHEHGAVKSR-----FKSITALEKIAQLEQVQDEAR 1842
 QY 1985 KVIALTERKMSLEQOINENVTTLKEGEGEKETYLQRPSSQOSSQWELRESLTKRKDQ 2044
 DB 1843 EKOQATKSLKQOKKILKEILLQVED-----ERKMAEQYKQOAEKGNANVAKLOKQ 1892
 QY 2045 LBEAEKESATNEIKNTAKISLEELIQQNSIINENAVSEENLRHSQOQVSELEOL 2104
 DB 1893 LBEAE-----ESORINNRRLQORE--LDEATESNEAGREVNALKSILR----- 1936
 QY 2105 SLTIKSRDHAFAQSKREKDEAVNKIASIABEIKILTKREMEFDSKES 2152
 DB 1937 ----RQNETSFVPSRSGRVIENMDGSEF--ETDTRDAD--FNGTKAS 1978
 RESULT 15
 US-09-815-242-12996
 ; Sequence 12996, Application US/09815242
 ; Patent No. US20020061569A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Karl L.
 ; APPLICANT: Zyskind, Judith W.
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Travick, John D.
 ; APPLICANT: Carr, Grant J.
 ; APPLICANT: Yamamoto, Robert T.
 ; APPLICANT: Xu, H. Howard
 ; TITLE OF INVENTION: Identification of Essential Genes in
 ; TITLE OF INVENTION: Prokaryotes
 ; FILE REFERENCE: ELITRA.011A
 ; CURRENT APPLICATION NUMBER: US/09/815,242
 ; CURRENT FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 12996
LENGTH: 6281
TYPE: PR
ORGANISM: Staphylococcus aureus
US-09-815-242-12996

Query Match 4.98; Score 718; DB 10; Length 6281;
Best Local Similarly 18.5%; Pred. No. 2e-19;
Matches 623; Conservative 617; Mismatches 1277; Indels 854; Gaps 150;

1 MSEDVAVCVVRPLIOREGDQA-NLQWKAANNITSOVNGTSGFNFDRVFNHESHSO 59
DB 3054 IADDAKATAYNFTDADQAKO--OAYNTAVTNAENIISKANG-----GNATQ 3098
QY 60 IYOEIAVPIIRSALOGYNGTIFAYGOTSSTGKTYTMGTNSIIPQAIQOEVEFKIIEIP 119
DB 3099 AEVQAIKQVNAKQALNGN--ANVQAKKDEATLINSNDLN--QAKQALK--QVYQ 3151
QY 120 NREPLKRVSTWEIYNEVKDL-----LCDDRKK-----PLETREDENRNVY 161
DB 3152 NATVAVGVNNVY--QTAQELNNAMTOLKQIADKEQTKADGNFVNADPPKONAYNQAVA 3208
QY 162 VAD-----LTRELVMPVPEHVQIWK--GEKNRHGETKM--NDHSSRSHITRMIVESDR 214
DB 3209 KAEALITATPDPVVTPESETRALNKVTOAKNDLNGNTNLATAKQOVQAHIDQLEPLNQAO 3268
QY 215 NDPNTSENCDAVAVYSHNLVDLAGSERASGTAGEVRLKEGCNINRSLEFLLGVYKILS 274
DB 3269 RD-EYSQIQTQATLVPMVNAIQOA--ATLNDAMTQKQGI-----AN 3308
QY 275 DGAAGGTINRDSKLTIRLQNSLGGNAKTVIICITTPVSFDETLSTLOFASTAKHVNTP 334
DB 3309 KAOIKGSENYHD-----ADTDR-----QTA 3328
QY 335 HVENVLDEALIKRYREIILDKOLENLESSSEFKAQAMAKEHT-----OLIAEIK 387
DB 3329 YDNAVTAABEL-----KOTNPYDPPTIOALIKVNDTQNALNGNOKLADAK 3377
QY 388 QLHKEREDRIWHLTNIVVASSQESQODORVKKRRVTWAPGKIQNSLHSGVSDPMLSR 447
DB 3378 ODAKTTIGTIDLND-----AKQALITQVQAPDIATVNNVQNAQNLN-----NMTN 3427
QY 448 LPRGFSKAKFSDMPFPEIDDSCTEPPDDALSMDSGIDAEWNLASKV--THREK 505
DB 3428 LNNALQDKT-----ETLNSI--NETDADAKKDAYTAVASHAEILISKANGSNASQ 3476
QY 506 TSLHQSIMIDGQISDVQVHDS--SKENQLOLYPKDSGDMAEGRKASFKEITSLQOOL 562
DB 3477 TEVEQAMORVNEAKQALNGNDNVORAKDAKQV--TNANDLNQAKQKQVDA--AQT 3533
QY 563 QSKEEKKEIVQSEFLKIAELEQLSYKAKNLEWNTSREHSAEV--QTVKEVEYVR 620
DB 3534 VAVNNTIKQTAQDNLQAMTQKQIADKQ-----TKANGFVADDDKQANAYNNAVAHA 3588
QY 621 EMSVLGDSGYNASND-----LQSSVYDGKRLSSSHDECIEHRKMLEQKIVDLEFTEENL 676
DB 3589 EQLISGTPNANVDPQVQVAAQVNOAKGDLNGNH-----NLQVAKDMAN 3633
QY 677 KKSE-----NDKQSSFODPHESTIOCEALIAEKANALEELALRDNFDNIIILENITKR 731
DB 3634 TALDQLEPLNQOPKTAALKDQVSHAEVLTGVNAIKQNA-----DALNNM--GTLKQ 3662
QY 732 EI-----ADLESLIKENQETNEFE-----ILEKTOKEHEQLIHEI 768

DB 3683 QIQANSQVPSVDFTQADQDQQAAYNNAQAQIANGIPPIVPLPDI-----VTQAVTMT 3738
QY 769 GSKKLVENAEKYNONLEE--DETQTKLKEQ-----EIQALEKRRADNLOKQVRN 819
DB 3739 NOAKDALNGDEKLQAKQALANDTLTLDLNOPOHDALRNOINQAOLATATEQOTKQON 3798
QY 820 FDLVSMGDSSEKICEEITQIKOSLSDAEAY-----TRDAQECSFRLSNLEKEME 872
DB 3799 VNTAAS-----NLKQGIANKDPTVKASENYHDADDKQVATNAVSQAEGLITQ 3846
QY 873 DTSWYNQKKAASLFERQLETKESNYKKRMEADLOKELOSAFNEINLGL-----LA 925
DB 3847 TTNPLNPDDELTRALTO--VTDAKNGLNGEAKLATEKQNAKDAVSGHTLINDAQKQALK 3903
QY 926 GKVPKDLISVLEKKEVSEFSKOLEKALEE--KNALENEVTCISEKFL--PNEVECLKN 981
DB 3904 GQI--DOSPEIATVNVQVQATISLDQAMDQSLQALNDRAQTLADGNVYNADPPKONAYKQ 3961
QY 982 QISKASEIMLLKQEGHSASTISKQETIMQESQIQLDDEVHTQOSKY-----1032
DB 3962 AVAKA--EALINKQSGTN-----EVAQVVESTINEVNAKQALNGNDLANA 4006
QY 1033 -QOTEEOYLEKKKMD--DLFEKYIRNKSEAEDL--REMEMIKGTMESEVYKIAOTK 1085
DB 4007 KOAKQOQLANTLHINDAQKSPESQITQAPLVTVTTINOKAQTLDHAMELLRNSVADNQ 4066
QY 1086 HEL-BETTRDKEQLLHEKKYFQAMQTFPTPLPSDSLPSKVLGNSQDPLEINDYHL 1144
DB 4067 TTLASEDHD--ATAQRONDYNOAVTAANNI--INQTTSP-----TMNDDVAV-- 4110
QY 1145 IALATERNNIMVCEITERN--SLKEQV--IDL-----NTQLOSLQASIEKSDP--QK 1191
DB 4111 -GATTQVNTKVALDGDENLAQAQANRRIDOLDLHNNQAKQOLQOSITQSSDIAVNG 4169
QY 1192 PKODLEGEVRLLEMMELKQHL-----TDSQLS-----TEKLOLEN 1228
DB 4170 HKQTAESLNTMAGNLINAIADHQAVEGRGNTADTQKQVNTAVNBAAMINQOTQON 4229
QY 1229 LEVTEKLOTLOEKMNTIETNELOTFNEEDLKAEDSLKQDSEIETIEF-----1280
DB 4230 ANQTEVEQAI-----TKVQTTLOALNGDHN--LOVAKTNATQADALATSLNDPQ 4276
QY 1281 ---ODELRAA-----OELREKQOLVSPFOQLDLS-----VGISPP--N 1316
DB 4277 KTALEKQVTAATFLYTAHQIEQNNATLNOAMHGLRQSTQDAAATKANSKYINEDQPEQON 4356
QY 1317 HDA-----VANDEKVSIGE-----VNSLOSEM-----LGERDELOTSCALVS 1355
DB 4337 YDQAVQANNTIIEQATLNDNNAIQAATVNTTKAALHGDVKLQNDKDH-----AKQTVS 4392
QY 1356 EL-----ELLRAHVKEGEGNEITKGLNGLEKEIIGK--SESEVYK 1396
DB 4393 QLAHLNNAQKMEEDTLIDSETTRTAVKQDLTEQAALDQMLQASIDKQATRASAVY 4452
QY 1397 SMLENLKEDNNKLQEAQESKEQNOFSLEE--VFSSQKQVDETEVYKAOLKAAEERLE 1454
DB 4453 NAEPRKQSYDEAYONAESITAGLNPTINKGNVSSATQAYISSNALLDGYERLAQK-- 4510
QY 1455 IKDRDYELVQANTNLVEGLLETPLOADHEBDSIDRSEMEIKVLGEKLERNOYLLER 1514
DB 4511 -----QTAGSNLHDLQILPRAQOQALENGQNNATNGEV--AQKLEBAQALMQA 4557
QY 1515 LOEKLELSKLEILOKEMETSVLLKDDLOK--LESLSLENNITLKNIDITLKHSST 1571
DB 4558 ME-----ALRNSIQDQOQTEAGSKFIEDKPOKDAVQAAVQAKDLINQTNPTL-----D 4608
QY 1572 QAOLOKTOQOELAKNLALASDNCPTIQEKEETADCVHPEEKILLITELHDKQKTNEO 1631
DB 4609 KAOVQOLQAVNOAK-----DN--LHGDKLADDKQHAATD-----LNLQNLGNNOQR 4654
QY 1632 KLHE-----KNELEQAVQELKC--EVEHLKMSIEMSKSLSLSS--LOHEKHDEQO 1678

```

Db 4655 QALESOJNNATRGVNAQKLAFAKALDQAMQALNSIODQOQTESGSKFTNEDEKPOKDAY 4714
QY 1679 LILAKQOMQVVTQ-----EKKELQOQHHEHLEAVEHDLKNEIELGMLFKNEAQOXTKEQ 1732
Db 4715 QAAVQNAKDLINOTGNPTLKSVOVEQ---LTOAVTTAKNLEHODKLANDQOQAVTTVN 4770
QY 1733 CLINENKELESOHRLQCEIEELMKSLKDESALETTKESEOKYININQEMWLEME 1792
Db 4771 ALPLNNAQOQ-----ALTDALINAPTRTEVAOHV---QATATELDHMET 4812
QY 1793 LKNSQRTVIAERDQ-----LODDLRESVEMSIETODDLKRAQEAALQOQKRVQELISQIS 1847
Db 4813 LKKNVDVNTDKAOPNTYEASTDSTKKEAVDQALQAAESITPTGNSNANKDAVOV---LT 4869
QY 1848 VLEKISILE-NOMLYVANV-VKETTLESDDLN-----OSKOHU-----PSEIFTLSS 1892
Db 4870 KLOKEKELNGENENVAKQAKQAKOTIDQULHNLADQATKONIDQATKLOPIAEVLDOA 4929
QY 1893 LSLKEFEPALEQAEKRDARAKTIDITERISNIE-----EOLLOQATNLKETLY 1942
Db 4930 TOLNOSMDLOQAAVNEHAN-VEQTVDYTOADSPKONAYKQAIADAENVLKNANKQOQVDQ 4988
QY 1943 ERESLLOK-----EOLAL-NTEHLRETLS-----KDLALGKMEQERD-----EAA 1993
Db 4989 ALQOITNAKQALNGDERVALAKTNGKHIDQULNNAQODGFKGRIDSDNLMOIQOIV 5048
QY 1984 NKVIALTEKSSLEEQINENVTTLKEGEGEKEFPYLOPSPKQOSSOMEELRESLKTDL 2043
Db 5049 DEKALNRADQOLSOQETIDN-----EGRTKSGTNYV-----NADTOVKOV-----YDE 5091
QY 2044 QLEBAKEISEBATEINIKNLTAQ-ISSLEEBELLONASTINEAVSERULRSKQOLVSELE 2102
Db 5092 TVDKAKQALTKSTGO---NLTAQOYIKLNDAY---TAAKALNGEERLNRKKAALORLD 5145
QY 2103 QLS-LTLKSDHAF-----AOSKREKDEAVNIAKSLAEIKLILKEMDEFR----- 2147
Db 5146 QULHNLNNAQOLATQOINNAETLNAKASRAINRAKTLKDNAMGAVOQYIDEQHLGYISSTNY 5205
QY 2148 -----DSKESLOEQSSHISEEL-----CYKTELQMLKQOKEDINN-----KLAERVK 2190
Db 5206 INADDNLKANYDINAIAVAHELDKVOGNALAKAEBQLKONIIDQONALNGDQOLANAKD 5265
QY 2191 EVDDELLOHSSLEKQOLDQOIMELNEKLRNYE-LCEKMDIMEKEI-----SVLRIMON 2242
Db 5266 KANAFVNSLNGLNOQ-----OODLAHKAINNADVSDVTDIYVNOQIDLNDAMETLKLHVDN 5321
QY 2243 E-POQEEED-----DVAERF--DILESBNQEOIOLMEKISAVYSEBOHTLSS 2285
Db 5322 EIPNAEQTVYQONADNAKTNFDDAKRLANTLNSDNTYNVDINGALQAVNDAIHN----- 5377
QY 2286 LSSELOKETEAKHCHMLNIKESLSLTSRSFGLQTEHVKLNTOLQTLNKFVYVYRTAA 2345
Db 5378 -----LNGQORLODARKDAIQSIN-----QALANKLKEIEASNA 5411
QY 2346 VKEDHSLIKXEKDLA-----AEQRHDELRLQOLCLEBQGRKWSASASEILKFCIEF 2399
Db 5412 TDQDKLAKAKAEELANSIINNINKATSNQAESQVQTAGNH-----AIEOVHANEI-- 5462
QY 2400 LNELLEFKKANIOSVODFSEVQVFLNOV-----GSTLOEELEHKKGFOMLEBFGDLHYD 2455
Db 5463 -----PKAK-IDANKYVDQVQVALIDEIDRNPULTKEKQALKDRINOIQLQO-GHNGIN 5514
QY 2456 AKRLSEGMQOENRIASTIOLTKRLKAVVQSK--IOREITVYLNQEPARKLOEKKEONKE 2513
Db 5515 NAMTKKEIEBQAKQALQALODIKDLYKAKEDAKODVQVQVALIDEI-----DQNP 5566
QY 2514 LMRMEHKGSPASVMEENARLGLIKTVODESKLOSRIKML-----ENELN-LVKDD- 2566
Db 5567 L-----TDKEQALKDRINOIQLQOQHNDINNALTKEEI 5599
QY 2567 AMHKGCVAILQD-KILSRNAEALNMOYKLTFKODNLOA-----AMKEIENL 2614
Db 5600 EQAKQALQALQALODIKDLYKAKEDAKNAIKALANAKRQOINSNDPLTPEQAKALKEIDEA 5659

```

```

QY 2615 QKMVAKAVPYKEEIDNL-----KTKVYKIEMEKI--KYSKATDQEIAYLKSC 2660
Db 5660 EKRALDN-VENAOITIDQNLNGMLGDDIDINTHVWEVDEQPAVNEIFENRPEIILVNGEL 5718
QY 2661 LEDKEEGLR-----RLKEELRRADQNDTTCVCPKDYQAKASTPPYTCGGSGGIVOST 2712
Db 5719 IYHRDITTEQDILAHINLIDQLSAEVIDPSPYATISDSL--TAKVEVTLIDSSKVIYV 5776
QY 2713 AMVLOSE-----KALERELSHYKKYHHLSTMSSEDRK-----KTRAKGDHNS 2759
Db 5777 PVKVEKELSVRAQOALESTENAAQKINEINNSVTLTLEQKEAIAEYKRLQOALIDHY 5836
QY 2760 SHTGSSHRGSPHKETERYR-H-GVTPPERSEMPSLHLSPPKSESSTRVYSPNRSEIYSOL 2818
Db 5837 NDAPDVHVSVEIQOQBHAHIEQFNPEQFTIEQAKSNAKSIEDALQHM-----DELKART 5892
QY 2819 VMSPGKTGMKHILSPSKVGLHKRRALSPNRSEMPHOYVISPCKTGLKHLKLTSTLFDNL 2878
Db 5893 DLTDKE-----KQEAIA-KLNOLEKEQAI-----QAIQRAQSIDEI 5926
QY 2879 SSPCKOQVOENLNSPKGLFDVSKSMPCPSQOFPDNSKLGDF--SELNTAESNDKSOA 2936
Db 5927 SEQLEQPKAOKKANPTELAKEL-AKKQOEAISRKDSNEKINSIRNSEIGTAD--EKQAA 5983
QY 2937 ENMWYEAKKET 2947
Db 5984 MNQINEIYLET 5994

```

Search completed: May 23, 2003, 15:03:20
 Job time : 340 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 23, 2003, 14:51:55 ; Search time 305 Seconds
(Without alignments)

2970.240 Million cell updates/sec

Title: US-09-150-867-1

Perfect score: 14769

Sequence: 1 MSEGDAVKVCVRVRIPIQRE.....QAEWNYEAKKETAPBECKTS 2954

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 segs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL-frame+p2n.model -DEV=xip
-Q/cgn2_1/USFTO.spool/US09150867/rumat_21052003_165913_20985/app_query.fasta_1.3143
-DB=Issued.Patents.NA -OPMT=fastap -SUFFIX=rml -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS-human40.cdi
-LIST=45 -DOCALLIGN=200 -THR.SCORE=pcr -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USR=US09150867.@CGN.1.1.120.@rumat.21052003.165913.20985 -NCPU=6 -ICPU=3
-NO_XLPRY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSYBLCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: Issued.Patents.NA.*
2: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
3: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
5: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1153	7.8	10136	1 US-08-353-700-2	Sequence 2, Appl1
2	1153	7.8	10136	1 PCT-US95-16216-2	Sequence 2, Appl1
3	1061	7.2	8789	1 US-08-328-254-5	Sequence 5, Appl1
4	1061	7.1	4757	1 US-08-572-191-1	Sequence 1, Appl1
5	1051	7.1	4757	1 US-09-723-262-1	Sequence 1, Appl1
6	1051	7.1	4757	1 US-09-723-219-1	Sequence 1, Appl1
7	941	6.4	4127	1 US-09-592-054-7	Sequence 7, Appl1
8	915.5	6.2	4176	1 US-09-722-139-1	Sequence 1, Appl1
9	915.5	6.2	4176	1 US-09-721-832-1	Sequence 1, Appl1
10	915.5	6.2	4176	1 US-09-721-889-1	Sequence 1, Appl1
11	909.5	6.2	4308	1 US-09-592-054-1	Sequence 1, Appl1
12	867.5	5.9	3572	2 US-08-713-815A-2	Sequence 2, Appl1

13	866	5.9	30549	4 US-09-134-001C-322	Sequence 322, App
14	848	5.7	6773	4 US-09-166-350-27	Sequence 27, Appl
15	841.5	5.7	3837	4 US-09-724-517-1	Sequence 1, Appl1
16	841.5	5.7	3837	4 US-09-641-807A-1	Sequence 1, Appl1
17	841.5	5.7	3837	4 US-09-723-096-1	Sequence 1, Appl1
18	776	5.3	8948	4 US-09-643-597-119	Sequence 119, App
19	769.5	5.2	11091	4 US-09-134-001C-2243	Sequence 2243, Ap
20	759	5.1	3319	1 US-08-006-676B-2	Sequence 2, Appl
21	759	5.1	3319	1 US-08-282-845-1	Sequence 1, Appl1
22	759	5.1	3319	1 US-08-428-414A-4	Sequence 1, Appl1
23	759	5.1	3319	2 PCT-US94-00324-2	Sequence 2, Appl1
24	758.5	5.1	3709	4 US-09-541-782-7	Sequence 7, Appl1
25	758.5	5.1	3709	4 US-09-723-820-7	Sequence 7, Appl1
26	756.5	5.1	4911	4 US-09-718-692-1	Sequence 1, Appl1
27	756.5	5.1	4911	4 US-09-718-852-1	Sequence 1, Appl1
28	756.5	5.1	4911	4 US-09-718-815-1	Sequence 1, Appl1
29	754	5.1	13977	4 US-09-484-970B-60	Sequence 60, Appl
30	748.5	5.1	6306	1 US-08-195-487-3	Sequence 3, Appl1
31	748.5	5.1	6306	5 PCT-US93-06160-3	Sequence 3, Appl1
32	744.5	5.0	6306	1 US-08-466-390-3	Sequence 3, Appl1
33	744.5	5.0	6306	1 US-08-470-950-3	Sequence 3, Appl1
34	744.5	5.0	6306	1 US-08-467-781-3	Sequence 3, Appl1
35	744.5	5.0	6306	2 US-08-483-924-3	Sequence 3, Appl1
36	742	5.0	13307	3 US-08-836-022A-10	Sequence 10, Appl
37	742	5.0	13307	4 US-09-427-048A-10	Sequence 10, Appl
38	733	5.0	4223	4 US-09-541-782-5	Sequence 5, Appl1
39	733	5.0	4223	4 US-09-723-820-5	Sequence 5, Appl1
40	731	4.9	5093	1 US-08-468-036-23	Sequence 23, Appl
41	731	4.9	5093	2 US-08-376-843-23	Sequence 23, Appl
42	724	4.9	1421	4 US-09-592-054-5	Sequence 5, Appl1
43	724	4.9	1569	1 US-09-592-054-3	Sequence 3, Appl1
44	717.5	4.9	9551	1 US-08-056-200-93	Sequence 93, Appl
45	717.5	4.9	9551	2 US-08-800-644-93	Sequence 93, Appl

ALIGNMENTS

RESULT 1
US-08-353-700-2
Sequence 2, Appl1 application US/08353700
Patent No. 5599919
GENERAL INFORMATION:
APPLICANT: YEN, TIMOTHY J.
APPLICANT: RATNER, JEROME B.
TITLE OF INVENTION: NUCLEIC ACID ENCODING A
TITLE OF INVENTION: TRANSIENTLY-EXPRESSED KINETOCHORE PROTEIN,
TITLE OF INVENTION: AND METHODS OF USE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: DANN, DOERFMAN, HERRELL AND SKILLMAN
STREET: 1601 MARKET STREET, SUITE 720
CITY: PHILADELPHIA
STATE: PA
COUNTRY: USA
ZIP: 19103-2307
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/353,700
FILING DATE: 09-DEC-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: REED, JANET E.
REGISTRATION NUMBER: 36,252
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 563-4100
TELEFAX: (215) 563-4044
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 10136 base pairs

```

TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHEICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: HUMAN
US-08-353-700-2

Alignment Scores:
Pred. No.: 3,18e-83 Length: 10136
Score: 1153.00 Matches: 736
Percent Similarity: 37.84% Conservative: 581
Best Local Similarity: 21.15% Mismatches: 1163
Query Match: 7.81% Indels: 1001
Gaps: 144

US-09-150-867-1 (1-2954) x US-08-353-700-2 (1-10136)
QY 81 PheAlaTyrGlnThrSerSerGlyLysThrTyrThrMetGlyThrProAsnSer 100
Db 465 TTCACGAGAGACACTGATTCAGCAAA----- 494
QY 101 LeuGlyIleIleProGlnAlaIleGlnGluValPheLysIleIleGlnGluIleProAsn 120
Db 495 -----AAACAATAGAAAAAAGTGGAGAACGGAATTAAGG 530
QY 121 ArgGluPheLeuValSerTyrMetGluIleTyrAsnGluThrValLysAspLeu 140
Db 531 TGTAAATCTGAGCTTGAAGAACCAACAGCGCGAGCTGCAGATGCTCTGCAAT 590
QY 141 LeuGlyAspAspArgLys-----LysProLeuGluIleArgGluAspPheAsn 157
Db 591 CCATGGAAATCACACAAAAAATTTTACAACTCCACTAACCAAGTCAATATATTAGT 650
QY 158 ArgAsnValIleValAlaAspLeuThrGluGluLeuValMetValProGluHisValIle 177
Db 651 GGTTCACAGTAT---GAGATCTTAAGAAAGAAATATATATAAGAGTTGA----- 688
QY 178 GlnTrrIleLysGlyGluLysAsnArgHisTyrGlyLysIleThrLysMetAsnAspHis 197
Db 699 -----GACGAAAAAAGATTAAGACAGAGCTTAAGGCTTGACAGCT 740
QY 198 SerSerArgSerHisThrIlePheArgMetIleValGluSerArgAsp-----ArgAsn 215
Db 741 AAAAAAGCAGACGACTCTTCCACAGCCACCATGATCACCGCATTTGCCCGGCAT 800
QY 216 AspProThrAsnSerGlu---AsnCysAspGlyAlaValMetValSerHisLeuAsnLeu 234
Db 801 CAGGCTTCATCATCTGTGCTTCATGCGACACAGAGAAAGACCCCAAGTCAATCTTTCA 857
QY 235 ValAspLeuAlaGlySerGluArgAlaSerGlnThrGlyAlaGluGlyValArgLeuLys 254
Db 858 -----TCATATCTCTCAAGAACT----- 875
QY 255 GluGlyCysAsnIleAsnArgSerIlePheIleLeuGlnValIleLysLysLeuSer 274
Db 876 -----CCAATTAAGAGAGATTCTCT 896
QY 275 AspGlyGlnAlaGlyIlePheIleAsnTyrArgAspSerLysLeuThr----- 290
Db 897 GCATCTTACTTCTTGCGGAGACTAGAGGTACCTCCAAAGTGCATCACTTGCAAAATAGG 956
QY 291 ---ArgIleLeuGlnAsnSerIleGlyLysAlaLysThrValIleIleCysThrIle 309
Db 957 AAAAGAGATGCTAATAGCACTTCTTGCAATTTAGC----- 995
QY 310 ThrProValSerPheAspGluThrLeuSerThrLeuGlnPheAlaSerThrAlaLysHis 329
Db 996 AGTCCTCATCTTTGGAT-----CAATTAAGAGCGGAATCAAGAG 1037
QY 330 ValArgAsnThrProHisValAsnGluValLeuAspAspGluAlaLeuLeuLysArgTyr 349

```

```

Db 1038 CTAGAAAC-----AAGATTAATGACTTG-----GAATACCGCTCGCAAGACAT 1082
QY 350 ArgLysGluIle-----LeuAspLeuLysGlnLeuGlu--- 361
Db 1083 GAAAAAGAAATGAAGGCCAAGTGAATTAAGTTCAAGAACTCCAACTCCAACTGAGAAA 1142
QY 362 ---AsnLeuGluSerSerGluThrLysAlaGlnAlaMetAlaLysGluGluHisThr 380
Db 1143 GCAAAAGTGAATTAATTTGAAAAAGAAAGTTTGACAAATGTGTAGGATGAAGT 1202
QY 381 GlnLeuLeuAlaGluIleLysGlnLeu-----HisLysGluArgLysAspArgIle 397
Db 1203 AGAACACACACAAATACAGCCAGCGGTCAACCAAGTATCTCATTTGAAACAAACAG 1262
QY 398 TrpHisLeuThrHisIleValAlaLysSerGlnGluSerGln----- 413
Db 1263 AAAAAATGACGAGATTTGAGTTGTCAGCAGCAAAATGCAAAAGTCCAGATGTCT 1322
QY 414 ---AspGlnArgValLysArgValThrTrrAlaPro---GlyLysIleGln 431
Db 1323 CTGGAACAGAAAAATTAAGAAAAAAGAGCTTTCAGAGAGCTCTCCCTCAACAG 1382
QY 432 AsnSerLeuHisAlaSerGlyValSerAspPheAspMetLeuSerArgLeu-ProGlyAs 451
Db 1383 CGTTCTTCCAAACACTGACACGAGAGTGCATGCATGAGCCAGATCCACCCAGAG 1442
QY 451 nPheSerLysLysAlaLysPheSerAspMetProSerPheProGluIleAspAspSer 471
Db 1443 TTACAGCAAG-----CCAAGAAAT 1460
QY 471 LysThrGluPheSerAspPheAspAlaLeuSerMetLeuAspSerAsnLysIleAs 491
Db 1461 ATGCACACAGCTCTGACGCTGATGAATTAACATCACTAAGAACACAGCTAGAA 1520
QY 491 PAlaGluTrrPAsnLeuAlaSerLysValThrHisArgGluLysThrSerLeuHisGlnSe 511
Db 1521 AACAAATTTGAAAGATTAAAGCAAAAGTGTGC-AGAGCTGAACAGCGCTTCCAGCGAG 1579
QY 511 rMetIleAspPheGlyGlnIleSerAspSerValGlnPheHisAspSerSerLysGluAs 531
Db 1580 TCAGATCAAGAAAGACGCTGAGAGAGCAATGAGAG-----GAATGAAGAAAGAAA 1633
QY 531 nGlnLeuGlnTrrLeuProLysAspSerGlyAspMetAlaGluCysArgLysAlaSerP 551
Db 1634 CAACCTCTTAAGAGTCACTTGACAAAGGCCAGAGAAAGTGTCCACCTGAGAGCAGA 1693
QY 551 eGluLysGluIleThrSerLeuGlnGlnLeuGlnSerLysGluGluLysLysGlu 571
Db 1694 ACTCAAGAACATCAAAACAGTGTAAATCAGAGCCAGAAATTTGCAAGAAATGAAGAG 1753
QY 571 uLeuValGlnSerPheGluLeuLysIleAlaGluLeuGluGluGlnLeuSerValLysAl 591
Db 1754 GAAGAAATACCTTCAGAAACCATGTTAAGAGATCTTCAAGAAAAATA----- 1802
QY 591 aLysAsnLeuGluMetValThrAsnSerArgGlnHisSerIle----- 605
Db 1803 -----AATCAGCAAGAAACTCTTACTACTTGAAGAAACTGAA 1840
QY 606 -----AsnAlaGluValGlnThrAspValGluLysGluValAlaArgLysGlu 621
Db 1841 GCTTGCTGCGCTGAGTGTGAAGACGAGATGTGTCACAGACCTTTTGAAGAAAG 1900
QY 621 uMetSerValLeuGluAspSerGlyTrrAsnAlaSerAsnSerAspLeuGlnAspSer 641
Db 1901 A-----GAACATCACTGAAACCACTTAATGATTAAGTAAAGACAGAG 1945
QY 641 rValAspGlyLysArgLysSerSerHisAspGluCysIleGluHisArgLysMetLe 661
Db 1946 GAAGAGTCCAAAGCCCTTGCTGAGTCT-----TT 1975
QY 661 uGluGlnLysIleValAspLeuGluPheIleGluAsnLeuAsnLys----- 677

```

Db 1976 AGAGTAAAAAAGAAACAAATATGCAAGATTGAAAGAAAGAAACTCTGTTTCTGTG 2035
 Oy 678 -LysSerGluAsnAspLys-----GluYSerSerGluInaSpheMetGluSe 694
 Db 2036 GAAAGGCAAAACGAAACAACTTTTAACCTCAGATGCAATCAGAAAGAAACTTCAAG 20955
 Oy 694 r-----LleGluLeuYs----- 698
 Db 2096 TMAAATTATCAGCTTGGGAAACTGTGTGAAGACACAGCAAAATMAAAGTCATGATACAA 21555
 Oy 699 -----GluAlaLeuMetaLeuGluAlaAsnAlaLeuGluLeuAlaLeuMe 715
 Db 2156 CGAAGAGTAGAAGACCTGGAGAGTATGACACAGAGAAACCTAATGTGCGAGATCAGAAACT 22155
 Oy 715 TArgAspAsnPhenAspAsnAlleLeuGluAsnGluThrLeuYsArgGluIleAlaAs 735
 Db 2216 TCACAACTGTATAGACAGTAAGTACAGTGGAGGTAGAGACCCAGAAACTAGCTTATATGA 22755
 Oy 735 PleuGluArgSer----- 739
 Db 2276 GCTACACACAAAAGCTGAGTTCACAGATCAGAAACATCAGAAAGAAATAGAAATATGTG 23355
 Oy 740 -LeuYsGluAsnGluInaGluThrAsnGluPheGluIleLeuGluYsGluThrGluYsGlu 759
 Db 2336 TTTAAACATCTTCACAGCTTACTGGCAGAGTTGAAGATCA-----GA 2377
 Oy 759 uHtSGluAlaGluLeuIleHtSGluIleGluYSerLeuYsLeuValGluAsnAlaGlu 779
 Db 2378 ACACAACTTCAGTACTGTCA-----ATGAAATATATAGCAAAAGACCG 24222
 Oy 779 uMetCtyrAsnGluAsnLeuGluInaSpheuGluThrIlyThrIlySLeuLeuYsGluGlu 799
 Db 2423 GTGTTCAC--CAAGACTTGCATCCGAAATATGAGACCTCAGAGGATCTGTAAATCCAA 2479
 Oy 799 nGluIleGluLeu----- 803
 Db 2480 AGATGCTTCTGTGTACAAATATGAATCATCAGAGAGCTTTTGCTTTGATCAGCA 25399
 Oy 803 ----- 803
 Db 2540 GCTGCATGCATCATCTCTTTCGCAAAATATATATGAGAAACAAAGAACATGCCCTGCA 25999
 Oy 804 ----AlaGluLeuArgLysArgAlaAspAsn-----LeuGluYly 815
 Db 2600 GAGGAGTGAATGTCTGTATAGAACACCAAGATCCGAAAAATCTGCCATCTACACAAA 26599
 Oy 815 sLysValArgAsnPhenAspLeuSerValSerMet-----GlyAspSerGlu 830
 Db 2660 TAGAGTTGATCTACTGGAATTTTCATTTAGAGTCTCAAAAACGATGAACCTGACCTGCA 2719
 Oy 830 uLysLeuCYsGluGluIlePheGluLeuYsGluInaSerLeuSerAspAlaGluAlaValThr 850
 Db 2720 AAACCATGTGAAGAGTGTGTGCAAAATCAAAAGAGAAATAGAA-----GAAATCTCAT 2773
 Oy 850 TArgAspAlaGluYsGluYsSerPheLeuArgSerGluAsnLeuGluLeuYsGluYly 870
 Db 2774 GAAAGCAGAACACATGCATCAAAAGTTTGTGTGCTGAACCAAGTCAGCCATTAGTAAGTT 28333
 Oy 870 sMetGluAspThrSerAsnTrpTyrAsn-----Glu 880
 Db 2834 ACAGAGAGAACACTTGTGCTCACCGAGATGTGTGTGCTGAACCTTAAGTGCCTTGAGAA 28933
 Oy 880 nLysGluLysAlaLaserLeuPheGluYsGluLeuGluInaThrGluYsSerAsnTyrLy 900
 Db 2894 CAAGGAAAAAGAGCTGCACACTTTTAAATGATATAGGTAGAAACTGACGCGCAGAGATTCA 29333
 Oy 900 sLysMetGluY-----AlaAspLeuGluYsGluLeuGlu----- 911
 Db 2954 AGAATTTAAAAAAGCAACCATCTACTGTGAAGCTCTCTAAAGGAGGATCAAACTTTATTC 30133
 Oy 912 -----SerAl 913.
 Db 3014 CGAAACCTTAGCTTGAGAGAAAGAAAGAAATATGATTCATCATTTCTTTAAATAAAGGGA 30733

QY	913	apheasnclulleasnfytleuasnclyleu	-----	AlaGly	927
Db	3074	AATTGAAAGAGCTGGCCCAAGAGATGGGACTTTAAGAAATTAATGCATCTTAAATCA			3133
QY	927	svalProragrspleuuseratrgval			939
Db	3134	AGAGAGATGAACCTTAATCCAGAAAGATGAGACTTTTGCAAACTAATATGATGAAGGA			3193
QY	939	uLysValSerGluPheSerLySGlnLeuGln	-----	-----	950
Db	3194	GAAAGCACTTTCAGAGTTACTGTATCTGATCAAGCAAGAAAACTATTATTACTCAAG			3253
QY	951	-----	-----	-----	955
Db	3254	ATGTGACAAACCCGAAATGCATGTGAGACTTTACTCAAAATACAAAGCCAGCAGGA			3313
QY	955	uLysAsnAla-----	LeuGlnAsnGlnValThrCysLeuSerGluTyrLeuPhe	971	
Db	3314	AAAGACTTCTAAATTAGAAATGCTGCTGAATTAATGACACTAGCTTTGTGAAGAATAG			3371
QY	971	eLeuProAsnGlnValGluCysLeuYsaAsnGlnIleSerLySAlaSerGlnGlnIle	991		
Db	3372	---AAATATAGTTGGAACAGCTAAAGGAAGCAATTTGCAAGGAACACCAAGAAATCTT			3427
QY	991	Ileu-----	LeuYsGlnGlnGluGlnHisSerAlaSerIleIleSerLyGlnGlnIleI	1010	
Db	3428	AACAAATTTACACTTTTGTGTGAAGAAACAAATCAGATCTGATCTGTGATTTGGACACT			3487
QY	1010	eMetGln-----	-----	GluGlnSerGln-----	1016
Db	3488	GCAGCAAGCTGTGAGATCTGATGACAGATACCAAAACAAATTTCAACAGAGCTGG			3547
QY	1017	-----	GlnIleGlnIleuThrAspGlnValThrHisThrGlnSerLyVal	1032	
Db	3548	TGGTTTAAGCAAGAAATCATGACTTTAAAGGAAGAAACAAATTCAAAGAAAT			3607
QY	1032	1-----	GlnGlnThrGlnGlnGlnIleuTyrLeuGlnMetLeuLySmetHisAsp--	1047	
Db	3608	TAAATACCTTATTAACAAGATGACAGTCAAGCTGATGAAGATATGACATCAAACTGAATG			3667
QY	1048	---AspLeuPheGlnYsTyrIleIleArgAsnLySerGlnAlaGlnAspLeuLeuGln	1066		
Db	3668	TCAAAATGTAAACATCAGAACCAATTTAGAACTGTGAAAGAAAGAGAGTGAAGAA			3727
QY	1066	uMetGlnAsnLeuYsGlyThrMetGln-----	-----	-----	1075
Db	3728	TCAATGTAATTTTAAACCTGACATGATCTGTGAAGTTAAAGAAATTTCTGATAGATTA			3787
QY	1076	-----	SerValGlnValLySAlaAspThrLySHisGlnLeuGlnGln	1090	
Db	3788	TAAATCCGAGTGTGCAATTAAGACATGATGCTCAAGAAATTAAGCAATTTAAATCCAGA			3847
QY	1090	uThrIleArgAspLySGlnLeuLeuHisGlnLySlyTyrPhePheGlnAlaMetGln	1110		
Db	3848	AAGTAGAGAGAGAGAGAGTGCCTGCCAGCATGA			3886
QY	1110	nThrIlePheProIleThrProLeuSerAspSerLeuProProSerLySLeuValGlnGln	1130		
Db	3887	GACAACTT-----	AGAGAGAGATCTTTGAACACAGCAATTTTCAAGACAT	3928	
QY	1130	YasSerGlnAspProIleGlnIleAsnAspTyr-----	-----	-----	1144
Db	3929	GCAGTCACAAAGAAATTTAGTGCGCTTAAAGACTGTGAATATGATCGCGAAGAAAGTATAT			3988
QY	1142	-----	HisAsnLeuIleAlaLeuAlaThrGlnTyrAsnAsnIleMetValCysLeu	1158	
Db	3989	TTACAGGCTCATAGTTGTGCACAAAGTCAAAAGCAATGCACACCTTCAGTGCCTCT			4048
QY	1158	uGluThrGluArgAsnSerLeuYsGlnGlnValIleAspLeuAsnThrGlnLeuGlnSe	1178		
Db	4049	GCAAACCAATTAAGACAGCTGAATGAG-----	CTAGAGAAATATGTGAAT	4096	

QY 1178 rleuGlAlaInserIleu-----LysSerAspLeuG 1190
 |||||
 Db 4097 ACTGAGGCTGAAAAATGATGAACCTGTAAGTACAGCTGATGATTCAGAGTACAGATATGAT 4156
 QY 1190 nlyPProLysInAspleuGluGluGluValLysLeuLeuLeuGluMetGluLeu 1210
 Db 4157 CACAGCACTAGAAAAATGGCAAGAGCTAGGCAAACTACTAATGAACTTAAATAT 4216
 QY 1210 uLyGly----- 1212
 Db 4217 AAATGATGACAGTGTCTCTCCATGCTGAGTTAGTGGAGACATACCAGAGTGATTT 4276
 QY 1212 ----- 1212
 Db 4277 TGTGTAACAACCAATGATACAGACACCTGTGTCTTGGCTCCATTTGACAGAGTAATTC 4336
 QY 1213 -----HisLeuThrAspSerGlnLeuSerIleGluLysLeuGlnLeuGluAsnLeuG 1230
 |||||
 Db 4337 CTACGAGCACTTGACA-----TTGTCAAGACAAAGAGTTCAAATGCACTTGGCGA 4387
 QY 1230 uValThrGluLysLeuGlnThrLeuGlnGluGluMetLysAsnIleThrIleGluArgAs 1250
 |||||
 Db 4388 ATTCAGAGAGAAATTTCTTACTTACAAAGTGAAGACAAATTTTACATGATCAGCACTG 4447
 QY 1250 nGluLeuGlnThrAsnPhelLysLeuLysAlaGluHisAspSerLeuLysGlnAspLe 1270
 |||||
 Db 4448 TCAGATGAGCTCTAAATATGTCAGAGCTGCAAGCTATGTTGACTCATTAAG----- 4499
 QY 1270 uSerGluAsnIleGluInserIleGluThrGlnAspGluLeuArgLalaGlnGlu 1290
 |||||
 Db 4500 -GCCGAAATTTGGCTCTTGCA-----ACGAATCTGAGAACTTTCAAGGTGA 4546
 QY 1290 uLeuArgGluGlnLysGlnLeuValAspSerPheArgGlnLysLeuAspCysSerVa 1310
 |||||
 Db 4547 CTGTGTGAAGAGAGATGACGTG-----GGCTTGAGAGAGGGGCTGCTCCATCCCTGTC 4600
 QY 1310 lGlyIleSerProAsnHisAspAlaValAlaAsnGlnGluValSerLeuGluGly 1330
 |||||
 Db 4601 ATCTCTGTGTGCTGCTACAGCTCTACTGCTTAC-----AGTTGGGAGA 4645
 QY 1330 uValAsnSerLeuGlnSerGluMetLeuArgGluGluArgAspGluLeuGlnThrSerCy 1350
 |||||
 Db 4646 C-----TCCTCTTTTACAGACTCTTTAGAA-----CAGACAGAGAGA 4684
 QY 1350 sLyAlaLeuValSerGluLeu----- 1358
 |||||
 Db 4685 TATGCTCTTTGATGATTTAGAAAGGGCTGTTTCAGCAACCAAGTCAGTATGATGA 4744
 QY 1359 -----LeuLeuArgLalaHisValLysSerValGluGluLysAsnLeuGlu 1374
 |||||
 Db 4745 AGTATTTTGCAGCAGCTGTCAGACCTATGTTGACTGCTTAAAGGCCGAAATTTGGCTTT 4804
 QY 1374 ethrLysLysLeuAsnGlyLeuGluLysGluIleLeuGlyLysSerGlu-----G 1391
 |||||
 Db 4805 GTCAGAGAACTGAGAAATTTCAAGGTGACTGTCGTAAGAGATGACAGTGGGCTTGA 4864
 QY 1391 uSerGluValLeuLysSerMetLeuGluAsnLeuLysGluAspAsnAsnLysLeuS-- 1410
 |||||
 Db 4865 GGAAGGGCTGCTGCATCTGTCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCA 4924
 QY 1411 -----GluGlnAlaGluLysSer----- 1417
 |||||
 Db 4925 TTGGGAGACTCTCTCTTTTACAGACTCTTTTAAAGACACAGAGATGCTCTCTCTTTT 4984
 QY 1418 -----SerLysGlnAsnGlnPheSerLeuGluGluValPhe----- 1429
 |||||
 Db 4985 GAGTAAATTTAGAGGGGTTGTTTACAGCAACAGTCAGAGTGAAGTATTTTTCAG 5044
 QY 1430 -----SerGlySerGlnLysLeuVa 1436
 |||||
 Db 5045 CAGCTGCAAGAGAGAGAAATCTGACAGAAAGAACCCCTTGGCCCCAGCGGAGGGGTGT 5104
 QY 1436 lAspGluIle-----GluValLeuLysAlaGlnLeuLysAlaAlaGluLys 1452

Db 5105 TGAAGACTGAGTCCCTGTGTAGAGGTGACCGGCACTCCCTCGAGAGACTAGAAAGAA 5164
 |||||
 QY 1452 gLeuGlu-----IleLysAspArgAspTyrPheGluValGlnThrAlaAs 1468
 |||||
 Db 5165 AATGGAAGTCAAGGATTTGAAAAATTAAGAAATTCAGAACTCGACGCTGATTAATAG 5224
 QY 1468 nThrAsnLeuValGlu---GlyLysLeuGluThrProLeuGlnAlaAspIleGlu----- 1485
 |||||
 Db 5225 TTCTGAAGCAAGAGCTTACTGCTGAGTACAGAAACAGTATTTGTCAGAAATAGACAGT 5284
 QY 1486 -GluAspSerIleAspArgArgSerGluGluMetGluIleLysValLeuGlyLysLeu 1505
 |||||
 Db 5285 GCAACAGAACTGCAAGAGCTGACTGAGATGAGAGTCCAACTGGCGGCAAAAAAGAA 5344
 QY 1505 uGluArgAsnGlnThrLeuLeu-----ArgLeuGlnGluLysLeuGlu 1521
 |||||
 Db 5345 ACAGACGCAACACTGCTCACTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 5404
 QY 1521 uLeuSerAsnLysLeuGluIleLeuGlnLysGluMetGluThrSerValLeuLysAs 1541
 |||||
 Db 5405 CTTAAGTCTCG-----TCTTGGCTTGGCAT 5431
 QY 1541 pAspLeuGlnLysLeuGlnSerLeuLeuSerGluAsnIleLeuLysGluAsnI 1561
 |||||
 Db 5432 CGACACAGAAAGATGATTCATCAAGC-----CGAAATGAGAGCTG 5470
 QY 1561 eAspThrThrLeuLysHisSerAspThrGlnAlaGlnLeuGlnLys----- 1577
 |||||
 Db 5471 TGCAATATCAAAAGAACTTCAAGAACTCAGAAACTCAGAAACCAACAGATGATTTCA 5530
 QY 1578 -----ThrGlnGlnLeuGlnLeu-----AlaLys 1586
 |||||
 Db 5531 TCAGATTTGATGAAGATGCTGACAGAGACCTCATTCAGATTCAGAAATATATCA 5590
 QY 1586 sAsnLeuAlaIleAlaAspAspAsnProIleThrGlnGluLysLeuThr----- 1604
 |||||
 Db 5591 GACTGTGAGTGAAGAACCCAGAGAGTGTCTGGGAACTCCGCAATACCAATTA 5650
 QY 1605 -----SerAlaAspCysValHisProLeu----- 1612
 |||||
 Db 5651 TGAAGCTCCAGGGAGAGTAAACCCAGGGCTCTCAGATGCAATTTGCAATTTGCAT 5710
 QY 1613 -----GluGluLysIleLeu 1618
 |||||
 Db 5711 TTCTGCTCAATGCTTTGGTACTATGATTTCTGGGAAATCAGAGATATATCAATA 5770
 QY 1618 uLeuThrGluGlnLeuHisGlnLysThrAsnGlnGlnGluLysLeuHisGluLysAs 1638
 |||||
 Db 5771 TCTTCAAGTGGGGTAAAGAGCATCAATGAGAAATTTGAGATTACTTCATGTGTATGA 5830
 QY 1638 nGluLeuGlnAlaGlnValGluLeuLysCysGluValGlnHisLeuMetLysSer 1658
 |||||
 Db 5831 GGAAGCTGACAG-----AAAGTTGAAGTTCTCTAAATGAAT 5869
 QY 1658 tIleGluSerLysSerSerLeuGlnSerLeuGlnHisGluLysHisAspThrGluGln 1678
 |||||
 Db 5870 GAAAGAAATTAAGTCAAACTC-----CATTTACAGAGGTGA 5908
 QY 1678 nLeuLeuAla-----LeuLysGlnGlnMetGlnValValThrG 1691
 |||||
 Db 5909 ACTAATGACCAAAATTTGAAGCATGCAATGGAATTTGGAATTTGGGGAATTTAGAA 5968
 QY 1691 nGluLysGluLeuGlnGlnThrHisGlnHisLeuThrAlaGluValAspHisLeu 1711
 |||||
 Db 5969 AGAAACCTCAGATTTAGAGAAATTTGGAATATTTCTCTGATTCACAGAGATTTCT 6028
 QY 1711 sGluAsnIleGluLeu-----GlyLeuAsnPhelLysAsnGluAlaGln---GlyLysTh 1728
 |||||
 Db 6029 CCAGAGAGTGAAGCTTCTGAAGGCTCAATCTGATTTAGAAATGCATGACATTAATC 6088
 QY 1728 rThrLysGluGlnCysLeuLeuAsnGluAsnLys----- 1739

Db 6089 ATCAGCTGAAGATATGAGATTAATGTGCCAAGTGAATGACAGCTGGAAGAGAGATT 6148
 QY 1740 -----GluLeuGlnInSerGlnHisArgLeuGlnCysGlu----- 1751
 Db 6149 TCTTGAATGTGGAAATAGCTGAGTAGAGATCGAATCGAGAAAGCTAGCATTTGACATGA 6208
 QY 1752 -----TLeuGlnLeuMetLysSerLeuLysAspLysGlnSerLeuAlaLeuGlnTh 1768
 Db 6209 AGCCCTTACTGAGAGCTGACTTGAAGAGTAGTCAAAACAGAGAAAGATATTTAGAAAA 6268
 QY 1768 rLeuLysGlnSerGlnGlnLysValIle---AsnLeuasnGlnGlnMetGlnMetValme 1787
 Db 6269 AGACAAATGAATAATAGCAAGAGTTATTGTCTGAGCAAGACAGTCTCAGAGTGTAC 6328
 QY 1787 rLeuGlnMetGlnLeuLeuLysAsnSerGlnArgThrValIleAlaGlnLysArg----- 1804
 Db 6329 AAGTGAGAGAAACAGCTTGTGTGAGAAATTAAGATCTATGTCAAAAAAACACGAGCACT 6388
 QY 1805 -AspGlnLeuGlnAspAspLeuArgGlnSerValGlnMetSerIleGlnLys----- 1821
 Db 6389 GGATCACTGTGTGAAAAAATGAAGAGAAACACAA---GAGCTGTGCTCATCAAG 6445
 QY 1822 -----GlnAspAspLeuArgLysAlaGlnGlnLysAla 1832
 Db 6446 TGAAGTCTCCATTGCATTTCAGTGCAGAGCGCAGAGGTGAAGGAAAAAGCGAACTCCT 6505
 QY 1832 uGlnGlnGlnLysAspLysValGlnGlnLeuThrSerGlnIleSerValLeuGlnGlnLys 1852
 Db 6506 TCAGACTTTCTCTCTGATGTAGTGTGAGTGTGTTAAAGACAAACCTATCTCCAGAGAAA 6565
 QY 1852 sIleSerLeuLeuGlnLysasnGlnMetLeuLysAsnValAlaThrValLysGlnLysLeu 1872
 Db 6566 GCGCAGAGATTGGAAAAAGCATCAGAGCATGCTCTTTGCAAAATGTAGCGCGAAAA 6625
 QY 1872 rGlnArgAspAspLeuasnGlnSerLysGlnHisLeuPheSerGlnLeuGlnLysLeu 1892
 Db 6626 CCAAAATTGCACAACTGAATTAAGAGAAAGAAATTCCTGTCAAGAACTGAAACCTCGCA 6685
 QY 1892 rLeuSerLeuLysGlnLysGlnLysGlnPhe-----AlaLeuGlnGlnLys 1905
 Db 6686 GGGCAGAGCTAGTGAATACATTAATGAAGAACTGAATGTCTCCAGAGCTGTGGAGCGCG 6745
 QY 1905 agLysAspLysAlaAspAlaAlaArgLysThrIleAspIleThrGlnLysIleSerAs 1925
 Db 6746 ACTGATGAGAGAAAGGTAGTTCGATTGAGGCTGAGCTCAACAGAGAGAGAGCATCA 6805
 QY 1925 nIle-----GluGlnGlnLeuLeu 1931
 Db 6806 GCTGAGAGAGGATCGAGAAACAGAGATTGCGATTGAGGCCGATGAAGAAAGAGCAT 6865
 QY 1931 uGlnGlnAlaThrAsnLeuLysGlnLysGlnLysGlnLysGlnLysGlnLysGlnLys 1951
 Db 6866 GGCATATCGCAGAGAACTGAAGAGACCGGAGGAGAAATGATTCACCTT-----AA 6916
 QY 1951 sGlnGlnLeuAlaLeuasnThrGlnHisLeuArgGlnLysLeuLysSerLysAspLeuAl 1971
 Db 6917 GGAATTAAGATTGAGAACTTGAAGGAAATTCAGATGTCAAGAGAAACAGAGAGCAT 6976
 QY 1971 aLeuGlnLysMetGlnGlnLysGlnLysGlnLysGlnLysGlnLysGlnLysGlnLys 1991
 Db 6977 GATT-----CTTGATGCGCCAGAAATTCCAAGCA----- 7004
 QY 1991 uLysMetSerSerLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 2011
 Db 7005 -GAAATGAGAGACTTAACAAACACAAATGAAGAGAGGCGCAGAGAGCTGGAAGATTGGA 7063
 QY 2011 u-----GlyGlnLysGlnLysGlnLysGlnLysGlnLysGlnLysGlnLysGlnLys 2024
 Db 7064 ATTAAGACCTTGCACGTTAAGTGTGAAAAAGAAATCTGACAAACAAATTAACAGAAAA 7123
 QY 2024 sGlnGlnSerSerSerGlnMetGlnLysGlnLysGlnLysGlnLysGlnLysGlnLysGln 2044
 Db 7124 ACAAGGTCACTGTGACAACTAGACAAAGTTACTCTTCAATTAAGTCTGTAGAGAA 7183
 QY 2044 nLeuGlnLysGlnLysGlnLysSerGlnAlaThr----- 2056
 Db 7184 AAAGAGACACAGAGATACAGATCAAGAAATCTTAAGAACTCAGTGCAGATGCTTCA 7243
 QY 2057 -AsnGlnLysAsnLeuThrAlaLysIleSerSerLeu-----GluGlnGlnLysLeu 2074
 Db 7244 GAATCAGTTAAAGAGCTAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 7303
 QY 2074 uGln---AsnAlaSerIleLeuasnGlnLysValSerGlnArgGlnLysGlnLysGlnLys 2093
 Db 7304 GAAAGCCACAGACAGAGTGTAGACCCACCAATAGAGAGAGAGAGAGAGAGAGAGAGAG 7363
 QY 2093 rLysGlnGlnLeuValSerGlnLysGlnLysGlnLysGlnLysGlnLysGlnLysGlnLys 2108
 Db 7364 CATGMAAGCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 7423
 QY 2108 uLysSerArgAspHisAlaPheAlaGlnSerLysArgGlnLysAspGlnLysAlaValAsn 2128
 Db 7424 ACAACTGAG-----GAAAGTGACCATCATCATCATTTACTTAAGGTTAG 7468
 QY 2128 sIleAlaSerLeuAlaGlnGlnLysIleLeuThrLysGlnLysAsp----- 2144
 Db 7469 AGTGAGAGACCTTAAAGAGAGCTAGAGATGACAGAGCAAAACCAAGAGCATGACCTCT 7528
 QY 2145 -GlnPheArgAspSerLys-----GluSerLeuGlnGlnGlnSerSerHisLeuSe 2161
 Db 7529 TGAGGAGAGAAATTCAAAG 7588
 QY 2161 rGlnGlnLysCysThrLysThrGlnGlnGlnMetLeuLysGlnGlnGlnGlnGlnGln 2181
 Db 7589 CCAAAAGTGTGAGAGGCTGTGAATTAAGATGTTGTACTATTAATGATGAGAAAAAATCT 7648
 QY 2181 eAsnAsnLysLeuAlaGlnLysValLysGlnValAspGlnLeuGlnGlnHisLeuSerSe 2201
 Db 7649 GACAAATGAATTAACAAAG 7691
 QY 2201 rLeuLysGlnGlnLysAspGlnIleGlnMetGlnLysGlnGlnGlnLysLeuArgAsn 2221
 Db 7692 -----GAAATTAATTAATCATCATTTGAAAT-- 7718
 QY 2221 rGlnLeuCysGlnLysMetAspIleMetGlnLysGlnLysLeuArgLeuMetGln 2241
 Db 7719 -----ATTTTGAAGAAAAAGA 7725
 QY 2241 nasGlnPheGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 2261
 Db 7736 GCAAGAGAGAAAGTACAGATGAAGAAATATCAAGCATGCGCATGAGATGCTTCAAAACA-- 7793
 QY 2261 gAsnGlnLysLeuGlnLysMetGlnLysIleSerAlaValLysSerGlnGlnGlnGln 2281
 Db 7794 -----CAATTAAGAGCTCAATGAGAGAGTGCACCTGCAATTAATGCAAGAGAG 7846
 QY 2281 rLeuLeuSer-----SerLeuSerSerGlnLeuGlnLysGlnLysGlnLysGlnLys 2298
 Db 7847 CTGAAGGCCAAAGAGAGATCTTAGTACAGTAGAG----- 7866
 QY 2298 sHisCysMetLeuasnLysGlnSerLeuSerSerThrLeuSerArgSerPheGlySe 2318
 Db 7887 -----TGCTT-----GAACTTGAGAGAGCTCAGTGTCTACAGAGCTTGTATGA 7930
 QY 2318 rLeuGlnThrGlnLysValLysLeuAsnThrGlnLeuGlnLysGlnLysGlnLysGlnLys 2338
 Db 7931 GGGCAAAATTAATTAATTTGTTTTCGAATCTTCAAGTGAAGGCTCTCATCAAGAA----- 7965
 QY 2338 sValValLysArgThrAlaAlaValLysGlnAspHisSerLeuLysAspArgGlnLys 2358
 Db 7966 -----CTAGAGAGTGCAGCA 8002
 QY 2358 sAspLeuAlaAlaGlnGlnLysArgHisAspGlnLeuArgLeuGlnGlnGlnGlnGlnGln 2378
 Db 8003 GAAACTGAGAGAGAGATGAAGAAATCAGTAGACTGAAAAATCAATTAACAGAGCAAGA 8062

QY 2378 ugnhlslyarqlyserpseralaserdgluldeuylphe----- 2394
 DB 8063 GCAGCTTGCTCTAACTGTCCAGTGGAGAGAGCCAACTTGGAGAGCAAAA 8122
 QY 2395 -Cysgluilegphelheulnleuileupheylsvalaasnilleleinserva 2414
 DB 8123 CTTAGACAGCAAAATCTGACGTGGAATTCGAGCAAGATCCAACTCAATCCAA 8182
 QY 2414 lgnasparpheserclvalaInalpheleuansgnlvalglserthleuIngl 2434
 DB 8183 AAATGCCCTTTGCGAGGACACATTAGAGTGTCCAGAGTCTTCCAAAGATTTAGAGAA 8242
 QY 2434 ugnleuglu-----HislelysglypmetgIntrepleugluuphegly----- 2450
 DB 8243 TGAGCTTGATTCAGAAAATGSCAAAATGCTCTTGTGAAAAGTAACAAAATGAC 8302
 QY 2451 -----AspleuHisValaspaIalyslys-----le 2459
 DB 8303 TGCMAAGAACTGACCTGCAGAGAAATGCATGAGATGCAGCAAAAACAGCAGAGCT 8362
 QY 2459 usergluylmetgIngluInasnarqyllealaserthrlleuInleuInhrly 2479
 DB 8363 GCAGAGAGACCTGAGTGGAGAGAAAAATAGCTGAGAGAGTGCAGTACTGTGGA 8422
 QY 2479 sarqleuylsalaValalInserlysleleInarqyllethVallyleuansgl 2499
 DB 8423 AGAAATTAAGAGC-----AGCAAAAGATCAA----- 8447
 QY 2499 nphleglualeylsleuInglulysylsleuInasnllysgluInleuMetlarqymetgl 2519
 DB 8448 -----TTGAGAGAGCTCACACTAGAAATAGTGAATTAAGAGAG----- 8486
 QY 2519 uHshlslyProserAlaserValmetglugluInasnlalargleuInleuInlylele 2539
 DB 8487 -----AGCCTAGATTCATGCACAAAGACAGAGCTGGAAGAGAGAG----- 8528
 QY 2539 ulysrthValalInaspluserlysylsleuInserarqyleuMetleuIngl 2559
 DB 8529 -----AAAGTGAGAGAGAAATAGCTGAATATCAGCTACGG----- 8564
 QY 2559 uleuansleuVallyasparalameHlslysglylys-----Valalalele 2577
 DB 8565 -----CTTCATGAGCTGAAAACAAAACACCAGGCTTGTCT 8599
 QY 2577 ugnasplysleuUserarqsnalaglualagluleuansnlametglInalysle 2597
 DB 8600 TTGGAC-----ACAACAAACAGATGAGTAAGTAATCCAGACAGAGAAAT 8653
 QY 2597 uHrlylsylsInasparnsleuInalalameHlslysgluileuInleuInlyske 2617
 DB 8654 GACTTCTTAAGAGAAATGCTCTCAGTTCACAGAGCTGAGATGACCTTTTAAAGCTAG 8713
 QY 2617 tValalalysglYalValProtyrlysgluInleasparn----- 2631
 DB 8714 T-----AAAGAGAGCTCAATTAATTCATTGAACCTACTAC 8749
 QY 2632 -----leuylsThlyValalysylleuMetglulysylelystYr----- 2646
 DB 8750 TCAGATTTGGAGAAATGTAAGAAACCAAGATGAGCAATCTAAATATGTAATCAGTT 8809
 QY 2647 -----SerlysalalathrasplngluilleaIlyrleuylserys----- 2660
 DB 8810 GAAGAGAGAAATGAACTGCCAGGAGAAATGAAGTTGTTGTAATCCGTGAACA 8869
 QY 2661 -leugluasplysgluIngllyleuInarqylleuInleuInarqylaglnal 2680
 DB 8870 GCTGGAGAGGAGAAAGAGTACTGCAAGAAAGAACTCTCAACCTTCAGCTCAGAGGA 8929
 QY 2680 aaspaanarPthrThrValcysValProlysaasrtyrGluInlyalaserThrPheProva 2700
 DB 8930 GAAGCAGAAACA----- 8942
 QY 2700 lHrCysgllygllyserlylevalInserThralamleuValleuInsergl 2720

DB 8943 -----GTAAGTGTATGATACCAAGAGTGCATGAATTAACAACCTGA 8983
 QY 2720 u-----LysalalaleuIn-----ArgluInleuSerHlslyrly 2733
 DB 8984 GATCAAGAACTGAAAGAACTCTGAAAGAAAAAACCAAGAGAGGAGATGAATACTTGA 9043
 QY 2733 slsYrHshlsleuSerArgThrmetserSerGluInarqylslyrlyVal 2753
 DB 9044 TAGTACTGTCTCTG-----CTTATAGCCCATGAAAGTTAGAAAGCTTAAGA 9094
 QY 2753 alysseraspalanserSerHlsThrlyserSerHls-----Argcl 2768
 DB 9095 GATGTAGAGACACAAGTGGCCCATGTGTCTGCACAGCAATCTAAACAAGATTCGCCAGG 9154
 QY 2768 yserProHlslyThrcIuThrlyrghlslyProvalThrProgluInserclue 2788
 DB 9155 GTCTCT-----TTGCTAGTCTCAGTTCTCCAGGACCATCTCAAT 9196
 QY 2788 tProserleu-----HisleuGlySerProlylsyse 2799
 DB 9197 CCTTCTGTACTGAAGAAGGTATCATCTGGCCAAAATPAAGCTTCAGGCAAGAGCA 9256
 QY 2799 rgluserSer-----ThrlsaryalValserProasnarqylleu 2815
 DB 9257 AGATCCAGTGAATATGGAGAAATGTAGAGAGCCACACCTGTACCCAGAGAGCT 9316
 QY 2815 rser-----GluInleuValmetSer-----Proglu-LysThrlymeth 2828
 DB 9317 TTTTAAAAAAGCAAGAAAGACATCATGAGTGTATTCACCTCGCAAGAGACA----- 9369
 QY 2828 lslYshlsleuSerProserLysvalgllyleuHlslys-----LysArgalaleus 2846
 DB 9370 --CGAAGTACTAGAGTTTACAGCCAGAGAGCTTCAGAAAGTGTAAAGAGGTGTGC 9427
 QY 2846 erProasnarqyl-----sergluMetProHlnHlsV 2857
 DB 9428 TGACATCCCGACAGAAAGACTAGCCATATATCTGCAGAAAGAACACATGGCAA----- 9483
 QY 2857 allleSerProgluylsThrlyleuHlslys-----leuThrg 2871
 DB 9484 --CTGGACACAGCCCGCCGCTGCTGCACAGAAATTAAGGCTATCCCACTAGTGTCCG 9541
 QY 2871 luserThrleuPheasparnsleuSerProCyslysgluInlyValalngluansl 2891
 DB 9542 CAAGAGAAATC-----TTCCAGAGTCTCCAAACCAACAGCGTGTGGA-----G 9586
 QY 2891 euanserProlysglylsleuPheasp-----Vallyserl 2904
 DB 9587 CAGATCAAAAGGTCAAAAGTTCTCAGCGAGCCCAAGTAGATTCAGGCACATCTCCG 9646
 QY 2904 yserMetProtyrCysProserglu-----PhePheaspaanserlyleuGly 2921
 DB 9647 AGAAGCCACACAGAAATCGTCCAGTCAATATCTCTGAGAGAAAGTCCAGCTGACAG 9706
 QY 2921 spPheSergluInleuansThralagluserasnasplysergluIngluansrtr 2939
 DB 9707 CCCAGAGAGGGGCTGAGGCTCAAGCAGAGGCGCACTTGTCCAGCCCAAGAGCTGG 9762
 RESULT 2
 PCF-US95-16216-2
 ; Sequence 2, Application PC/TUS9516216
 ; GENERAL INFORMATION:
 ; APPLICANT: Yen, Timothy J.
 ; APPLICANT: Ratner, Jerome B.
 ; TITLE OF INVENTION: Nucleic Acid Encoding a Transiently
 ; TITLE OF INVENTION: Expressed Kinetochores Protein, and Methods of Use
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Dann, Dorfman, Herrell and Skillman
 ; STREET: 1601 Market Street Suite 720
 ; CITY: Philadelphia
 ; STATE: PA

```

COUNTRY: USA
ZIP: 19103-2307
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/16216
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,700
FILING DATE: 09-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Reed, Janet E.
REGISTRATION NUMBER: 36,252
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 563-4100
TELEFAX: (215) 563-4044
INFORMATION FOR SEQ. ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 10136 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHEICAL: NO
ANTI-SENSE: NO
PCT-US95-16216-2

```

Alignment Scores:

```

Pred. No.: 3,18e-83 Length: 10136
Score: 1153.00 Matches: 736
Percent Similarity: 37.84% Conservative: 581
Best Local Similarity: 21.15% Mismatches: 1163
Query Match: 7.81% Indels: 1001
Gaps: 144

```

US-09-150-867-1 (1-2954) x PCT-US95-16216-2 (1-10136)

```

QY 81 PheAlaTyrGlnThrSerSerGlyLysThrTyrThrMetMetGlyThrProAsnSer 100
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 465 TTCACAGAGAGACACATGATTCAGGCAAA----- 494

QY 101 LeuGlyIleLeuProGlnAlaIleGlnGluValPheLysIleIleGlnGluIleProAsn 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 495 -----AAACAATAGAAAAACTGGACACGAACTTAAAGG 530

QY 121 ArgGluPheLeuLeuArgValSerTyrMetGluIleTyrAsnGluThrValLysAspLeu 140
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 531 TGTAAATCTGAGCTTGAAAGAACCAACAGCTCGCGACATGCTCTCTCAAT 590

QY 141 LeuCysAspAspArgArgLys-----LysProLeuGluIleArgGluAspPheAsn 157
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 591 CCATGGAATAACACCAAAAAATTTTACAACTCCACTACCAACCAAGTCAATATTATAGT 650

QY 158 ArgAsnValTyrValAlaAspLeuThrGluGluLeuValMetValProGluHisValIle 177
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 651 GGTTCACAGATAT--GAAAGTCTTAAAGAAAAATATATATTAAGAGCTTGA----- 698

QY 178 GlnTrpIleLysGluGluLysAsnArgHisTyrGlyLysThrLysMetAsnAspHis 197
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 699 -----GACGAAAAAAGATTAGAGCAGAGGTTAAAGCTTCGACAGCT 740

QY 198 SerSerArgSerHisThrIlePheArgMetIleValGluSerArgAsp-----ArgAsn 215
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 741 AAAAAAGCAAGCAGACTCTTCACCAACCCACCATGATCAGCGCATTTGCCCGCAT 800

QY 216 AspProThrAsnSerGlu---AsnGlyAspGlyAlaValMetValSerHisIleAsnLeu 234
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 801 CAGCTTCATCATCTGTTCTCATGCGACCAAGAGAACCCCAAGTCAATCTTTTCA--- 857

```

```

QY 235 ValAspLeuAlaGlySerGluArgAlaSerGlnThrGlyAlaGluGlyValArgLeuLys 254
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 858 -----TCTAATTTCTCAAGAACT----- 875

QY 255 GluGlyCysAsnIleAsnArgSerLeuPheIleLeuGlnValIleLysLysLeuSer 274
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 876 -----CCAAATTAGAGAGATTTCTCT 896

QY 275 AspGlyGlnAlaGlyPheIleAsnTyrArgAspSerLysLeuThr----- 290
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 897 GCATCTTACTTTCTGGGAACTAGAGTGACTCCAAAGTCAATCAATTTTCCAAATAGG 956

QY 291 ---ArgIleLeuGlnAsnSerLeuGlyLysAsnAlaLysThrValIleIleCysThrIle 309
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 957 AAAAGAGATGTAATACAGCTTTCTTGGCAATCTTAC----- 995

QY 310 ThrProValSerPheAspGluThrLeuSerThrLeuGlnPheAlaSerThrAlaLysHis 329
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 996 AGTCCTCATCTTTTGAT-----CAATTAAAGCGCAGAAATCAAGAG 1037

QY 330 ValArgAsnThrProHisValAsnGluValLeuAspArgGlyAlaLeuLysArgTyr 349
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1038 CTAAAGAAAC-----AGATTAAAGAGTTG-----GAACTACGCTGCAAGGACAT 1082

QY 350 ArgLysGluIle-----LeuAspLeuLysGlnLeuGlu--- 361
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1083 GAAAAAGAAATGAAGGCCCAAGTCAATTAAGTTTCAAGACTCCCACTCCAGTGGAGAA 1142

QY 362 ---AsnLeuGluSerSerSerGlyThrLysAlaGlnAlaMetAlaLysGluLysThr 380
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1143 GCAAAAGTGAATTAATTAAGAAAAAGAAAGAAAGTTTGAACAAATGATGAGATGACTAGT 1202

QY 381 GlnLeuLeuAlaGluIleLysGlnLeu-----HisLysGluAlaGluLysArgGly 397
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1203 AGAACCAACGACATATACGACAGCGCTCAACCAAGATACTGATTTGACAAACAACTG 1262

QY 398 TrpHisLeuThrAsnIleValAlaLysSerSerGlnGluSerGlnGln----- 413
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1263 AAAAAATTCAGCGAAGATTGAGTTGTCAGCGACAAATGCAAGAAAGTCCAGATGTTCT 1322

QY 414 ---AspGlnArgValLysArgLysArgValThrTrpAlaPro---GlyLysIleGln 431
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1323 CTGACACAGAAATTAAGAAAAAGAAAGAGTTTCAAGAGAGCTCTCCGTCACACAG 1382

QY 432 AsnSerLeuHisAlaSerGlyValSerAspPheAspMetLeuSerArgLeu-ProGlyAs 451
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1383 CGTCTTCCAAACACTGACGACGACGAGTGCATCCAGATGAAGGCCAGACTCACCAGAG 1442

QY 451 nPheSerLysLysAlaLysPheSerAspMetProSerPheProGluIleAspAspSerVa 471
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1443 TTACAGCAG-----CCAAAGAT 1460

QY 471 LysThrGluPheSerAspPheAspArgAlaLeuSerMetLysAspSerAsnGlyIleAs 491
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1461 ATGCACACAGCTCTGACGCTGATTAACCTGATTAACCTGATTAACCTGATTAACCTGAT 1520

QY 491 PAlaGluTrpAsnLeuAlaSerLysValThrHisArgGluLysThrSerLeuHisGlnSe 511
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1521 AACAAATTGGAAGAGTTTAAACAAAGTTGTC-AGAGCTGAAAGCGCGTTCACAGCGGAG 1579

QY 511 IMetIleAspPheGlyGlnIleSerAspSerValGlnPheHisAspSerSerLysGluAs 531
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1580 TCAATGACAGAGAAATAGCTGAGAGAACATGAG-----CAATTAAGAAAGAGAAA 1633

QY 531 nGlnLeuGlnTyrLeuProLysAspSerGlyAspMetAlaGluCysArgLysAlaSerPh 551
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1634 CAACCTCTTAAGAGTCACTCTGAGCAAAAGGCCAGAGAAAGTGGCACTGAGAGGAGA 1693

QY 551 eGluLysGluIleThrSerLeuGlnGlnIleuGlnLysGluGluGluLysGlu 571
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1694 ACTCAAGAAATCAACAAAGCTTTAAATCAAGGCCAAATTTTGCAGAAATTAAGAAAGC 1753

QY 571 uLeuValGlnSerPheGluLeuLysIleAlaGluGluGluGlnLeuSerValLysAl 591

```


OY	1090	uThrIleargAspIysgluInleuLeuhtSgluYsYrYrPhepneGlnAlaMeGl	1110
Db	3848	AAGGAGAAAGAGAGAAAGAGCGCTGCAGATGAA-----TTACA	3886
OY	1110	nThrIlePheProIleThrProLeuSerAspSerLeuProProSerLysLeuValGluI	1130
Db	3887	GACAAAT-----AGAGAGACTCTGAACACACCAATTTGGCAAGACAT	3928
OY	1130	YAsnSerGlnAspProIleGluIleAsnAspYr-----	1141
Db	3929	GCACCTACCAAGAAATAGTGGCCTTTAAACACTGTGAAATAGATGCGGAGAAAGATATAT	3988
OY	1142	-----HisnleuIleAlaLeuAlaThrGluArgAsnAsnIleMetValCysLe	1158
Db	3989	TTTCAGGCGCTCATGATGTTGTTCACCAAGTCAAAACGCAATGCACACCTTCAGTGTCTCT	4048
OY	1158	uGluThrGluArgAsnSerLeuYsGluInValIleAspleuAsnThrGlnLeuGlnSe	1178
Db	4049	GCAACCAACCAATGACCAACGCTGAATGAG-----CTAGAGAAATAATGTGAAAT	4096
OY	1178	rLeuGlnAlaGlnSerIleGlu-----LysSerAspLeuG	1190
Db	4097	ACTCCAGCGCTGAAAAGTATGATCTGTAACTGAGCTGAATGATTCAGGTGCAATGTAT	4156
OY	1190	nLysProLysGlnAspLeuGluInGluIValIlyLysLeuLeuLeuGluMetGluLeu	1210
Db	4157	CACGACCAACTAGAGAAATGGCCAAAGAGTAGGGAACACTCAATTAAGATTAAATTT	4216
OY	1210	uLysGly-----	1212
Db	4217	AAATGATGACAGTGGCTTCTCCATGATGATTAAGTGAAGACATACCAGAGGTGAAT	4276
OY	1212	-----	1212
Db	4277	TGGGACCAACCAATGACAGACACCCGTGTCTTGGCTCCATTTGGACGAGAGATATC	4336
OY	1213	-----HisleuThrAspSerGlnLeuSerIleGluYsLeuGlnLeuGlnAsnLeuG	1230
Db	4337	CTACGACCACTTGACA-----TTGTCCAGCAACAAGATTCAAAGCACTTTGGCCA	4387
OY	1230	uValThrGlnIleLeuGlnIleThrleuGlnGluMetLysAsnIleThrIleGluArgAs	1250
Db	4388	ATTCAAGAGAAATTTCTTATCTTTACAAAGTGACCAAAATTTTACATGATCGCACCTG	4447
OY	1250	nGluLeuGlnThrAsnPheGlnAspLeuYsAlaGlnHisAspSerLeuYsGlnAspLe	1270
Db	4448	TCAGATGAGCTCAAAATGTCAAGAGCTGCAGAACCTATGTCATTAATAAG-----	4499
OY	1270	uSerGlnAsnIleGluInSerIleGluThrGlnAspGluLeuArgAlaAlaGlnGluG	1290
Db	4500	-GCCGAAATTTGGCTTGTCA-----ACGAACTCGAAGAACTTCCAAAGTGA	4546
OY	1290	uLeuArgGlnGlnLysGlnLeuValAspSerPheArgGlnGlnLeuLeuAspYsSerVa	1310
Db	4547	CTTGGTAAAGAGATGCAGCTG-----GGCTTGAAGAGAGGGGCTCGTTCATCCCTGTC	4600
OY	1310	IcGlyLeSerSerProAsnHisAspAlaValAlaAsnGlnIuYsValSerLeuGlyG	1330
Db	4601	AYCTCTTGTGTCTGCAGACGCTCTGATCTTGAAC-----AGTTGGGAGA	4645
OY	1330	uValAsnSerLeuGlnSerIleuMetLeuArgGlyGluArgAspGluLeuGlnThrSerCy	1350
Db	4646	C-----TCTCTCTTTTACAGACCTTTTAAAG-----CAGACGAGAGA	4684
OY	1350	sLysAlaLeuValSerIleuGlu-----	1358
Db	4685	TATGTCTCTTTGATTAATTGAAGGGGCTGTTCCAGAAACGAGTGCAGATGATGA	4744
OY	1359	-----LeuLeuArgAlaHisValLysSerValGluGlyGluAsnLeuGluI	1374
Db	4745	AGATATTTTGGACAGAGCTGTGCAGACCTATGTGACATTAAGAGGCGCAAAATTTGGCTTT	4804

QY	1374	ethrltyslsyeuansnclyleuenglulysgluileleuglyltysserglu-----gl	1391
Db	4805	gtcaacgagatgtggaaacattttcaagctgcttggtggaaagagatgacgtggccttggaa	4864
QY	1391	userglvalleuyssermetleugluasneleuylsgluaspaanasytleuyls--	1410
Db	4865	ggaggcgctcttccatccctgtcattcccttctgtgacctgacactctagctttagacg	4924
QY	1411	-----gluglnalagluglutyrsr----	1417
Db	4925	tttggagactctctttttacagagctctttttagaacagacagatattgtcttcttt	4984
QY	1418	-----serlsgluanglnpheserleugluvalphe-----	1429
Db	4985	gagtaattttagaaagggtgttttcagcaaacagctgactgtatgtatgaattttttgcag	5044
QY	1430	-----serglyserglntylsleua	1436
Db	5045	cagttcgacgagagaaatctgacacagaaaaaccccttggcccccagcgaaaggctgt	5104
QY	1436	laspelutle-----gluvalleuylsalaglnleuylsalalaglugluar	1452
Db	5105	tgaagagcttgagtcctctgtgaggtgacccggcagctccctcgaaagactgaaagaga	5164
QY	1452	gluelu-----llelysapartgsplyrphiegluleuvalglntthalas	1468
Db	5165	aatgcamaatgcagagattttgaaaaattgaaattcaagcactgacagctattttatmg	5224
QY	1468	nthrasnleuvalglu---glylsleuengluthtproleuglnalasphtasglu-----	1485
Db	5225	ttctgaaagcgaaagacttgacgtccctttaggaagcagttttgtcgaataatgaaacgtg	5284
QY	1486	gluaspserlleaspartgatsergluglumetglutltylsvalleuglyglutylsle	1505
Db	5285	gcacacagagactgcacagcgctgactgtgagatgagatgcacagttggccggcagaaagaa	5344
QY	1505	ugluatrgasnclntlytleuenglul-----argleuglngluylusleugl	1521
Db	5345	acagcgcgacacactgtctacttgagctggaagatgacacacttcacagagcttgcga	5404
QY	1521	uuserasnlystleugluileleuglnglysmetgluthtsrvalleuleuyls	1541
Db	5405	ctttaaattctctgg-----tctttgctttggcat	5431
QY	1541	paspserlnglntysleugluserleuensergluasneileuleuylsgluasnll	1561
Db	5432	cgacacagaaagatctatctcaagcc-----cgaaatgaagactg	5470
QY	1561	easpthrthleuylshsthsersapthrglnalaglnleuglnlys-----	1577
Db	5471	tgacattatcaaaamaacatcttcacagaaactacagaaamaacacamaacatgatgttca	5530
QY	1578	-----thrglnglngluenleu-----alaly	1586
Db	5531	tcagattttgtatpaaagactctagcagcagctcaatctpagaacttgcagaaatpactga	5590
QY	1586	sasnleuallellealalasersapancyproliethrglnglntylsgluthtsr-----	1604
Db	5591	gactgctgactgaaaacccacagagagctctctgggcaaacagctccacagatnaccatpaa	5650
QY	1605	-----seralaaspyvalhnsproleu-----	1612
Db	5651	tgagccttcacggagaaatpaaacccagggcctcttcagaaatgcattttctgaattgtcat	5710
QY	1613	-----gluglulyslleulele	1618
Db	5711	tttctggtctaatgcttttggtaacctatgcattttcctggggaattcgagaaatpaccatpaa	5770
QY	1618	uleuethrglngluenltystrhsanglnglnglntylsleuethlsglutyls	1638
Db	5771	ttcttcaactgcgggttaaaagacatccaaattggaattttgaaatttactttcatgtatgaa	5830
QY	1638	ncluleuenglinalaglnvalgluleuylsycsgluvalglunlthleuetyserme	1658

QY 2281 rleuLeuSer-----SerLeuSerSerGluLeuGlnThrGlnAlaHisIty 2298
 Db 7847 CTGTAGGCGAAGACAGCAATCTTAGTCAGTAGAG----- 7886
 QY 2298 sHisCysMetLeuAsnIleLysGluSerLeuSerSerThrLeuSerArgSerPheGlySe 2318
 Db 7887 ----TGCTCT-----GAACTTGAGAAAGCTCAGTTCGACAAAGCGCTTGATGA 7930
 QY 2318 rLeuGlnThrGlnHisValLysLeuAsnThrGlnLeuGlnThrLeuLeuAsnLysPheLys 2338
 Db 7931 GCGCAAAATATATATATATGTTTTCGAACTTCAGTGAAGCGCTCATCAAGAA----- 7985
 QY 2338 sValValItyrArgThrAlaAlaValLysGluAsnPhisSerLeuIleLysAspTyrGluLys 2358
 Db 7986 -----GTAGAAAGATGGCAAGCA 8002
 QY 2358 sAspLeuAlaAlaGlnGlnLysArgHisAspGluLeuArgLeuGlnLeuGlnCysLeuGln 2378
 Db 8003 GAAACTGGAGAGAGAGATGAGAAATCAGTAGACTGAAATTCAAATTCAGACACAGA 8062
 QY 2378 uGlnHisGlyArgLysTyrPheArgSerAlaSerGluLeuLysPhe----- 2394
 Db 8063 GCAGCTGTCTCTAAACTGTCCAGTGGAGAGACCAACTTGGAGAGACAAA 8122
 QY 2395 -CysGluIleGluPheLeuAsnGluLeuPheLysLysAlaAsnIleIleGlnSerVa 2414
 Db 8123 CTTAGAACTGAGAAATCTGACGCGTGAATTCGACGAGAGATCCAAAGTCTCAATCCAA 8182
 QY 2414 IGlAspAspPheSerGluValGlnValPheLeuAsnGlnValGlySerThrLeuGlnGln 2434
 Db 8183 AAATGCTCTTTCAGACACATAGAGTGTGCAAGATCTTCAAGAAATCTAGAGA 8242
 QY 2434 uGluLeuGlu-----HisLysLysGlyPheMetGlnTrpLeuGlnGluPheGly----- 2450
 Db 8243 TGAGCTGAATTCACAAAATGACAAAATGCTTGTGTGAAAAGTAAACAAAATGAC 8302
 QY 2451 -----AspLeuHisValAspAlaLysLys-----Le 2459
 Db 8303 TGCAAGAGAACTGACGTCGACAGAGGAATGATGATGCGACAGAAAACAGACAGCT 8362
 QY 2459 uSerGlnGluMetGlnGlnLysAsnArgArgIleAlaSerThrIleGlnLeuLeuThrLys 2479
 Db 8363 GCAGAGAACTGACAGTGGAGAAAATAGCTGACTGAGAGTTCGACTACTGTGGA 8422
 QY 2479 sArgLeuLysAlaValValGlnSerLysIleGlnArgGluIleThrValTyrLeuAsnGln 2499
 Db 8423 AGAAATTAAGAGC-----AGCAAGATCAA----- 8447
 QY 2499 nPheGlnAlaLysLeuGlnGlnLysLysGlnAsnLysGluLeuMetArgArgMetGln 2519
 Db 8448 -----TTGAAGGAGCTCACACTAGAAAATAGTGAATTGAAGAAC----- 8486
 QY 2519 uHisHisGlyProSerAlaSerValMetGlnGlnLysAsnAlaArgLeuLeuGlyIleLe 2539
 Db 8487 -----AGCTAGATTCATGCATGCACAAAGACAGCGTGAAGAAAGAGG----- 8528
 QY 2539 uLysThrValGlnAspGluSerLysLysLeuGlnIleSerArgIleLysMetLeuGlnAsnGln 2559
 Db 8529 -----AAACTGAGAGAGGAATAGCTGAATATCAGCTACGG----- 8564
 QY 2559 uLeuAsnLeuValLysAspAspAlaMetHisLysGlyLysLys-----ValAlaIleLe 2577
 Db 8565 -----CTTCATGAGACCTGAAAAGAAAGAACACACGCGCTTGCT 8599
 QY 2577 uGlnAspLysLeuLeuSerArgAsnAlaGluAlaGluLeuAsnAlaMetGlnValLysLe 2597
 Db 8600 TTTTGAC-----ACAACAAACAGTATGAGATGAATCCAGACATACCGAGAAATTT 8653
 QY 2597 uThrLysLysGlnAspAsnLeuGlnAlaAlaMetLysGluIleGlnAsnLeuLysLysKe 2617
 Db 8654 GACTTCTTAAGAAAGATGCTCAGTTCACAGAAAGCTGAGATAGACTTTTAAAGCTTAG 8713

QY 2617 tValAlaLysGlyAlaValProTyrLysGluGlnIleAspAsn----- 2631
 Db 8714 T-----AAAGAGAGCTCAATTAATCTTCAAAAGCTACTAG 8749
 QY 2632 ----LeuLysThrLysValValLysIleGluMetGlnLysIleLysTyr----- 2646
 Db 8750 TCAGATTTTGGAAGATGAGAAATCAAGAAATGACCAATCTAAATATGTAATACAGTT 8809
 QY 2647 -----SerLysAlaThrAspGlnGlnIleAlaTyrLysLysSerCys----- 2660
 Db 8810 GAAGAGAAATTAAGAGCGCCAGGGAATTAAGATTTGTTGATCAATTAATCTGTAACA 8869
 QY 2661 -LeuGlnAspLysGluGlnGlyLeuArgArgLeuLysGluGlnLeuArgAlaGlnAl 2680
 Db 8870 GCTGAGAGAGAAAGATACATCTGCAGAAAGAACTCTCAATTCAGAGCTGCACAGGA 8929
 QY 2680 aAspAsnAspThrThrValLysValProLysAspTyrGlnLysAlaSerThrPheProVa 2700
 Db 8930 GAAGCAGAAACA----- 8942
 QY 2700 IThrCysGlyGlyGlySerGlyIleValGlnSerThrAlaMetLeuValLeuGlnSerGln 2720
 Db 8943 -----GGTACTGTTATGATACCAAGTGCATGAATTAACAACCTGA 8983
 QY 2720 u-----LysAlaAlaLeuGln-----ArgGluLeuSerHisTyrLysLys 2733
 Db 8984 GATCAAAAGACTGAAAGAACTCTGAAAGAAACCAAGAGAGCGCATGAATACTTGGCA 9043
 QY 2733 sLysTyrHisHisLeuSerArgThrMetSerSerSerGluAspArgLysValThrLysAl 2753
 Db 9044 TAAAGTACTGCTCTTG-----CTTAAAGCCATGAAGAAAGTTAGAAAGCTMAAGA 9094
 QY 2753 aLysSerAspAlaHisSerSerHisThrLysSerHis-----ArgGln 2768
 Db 9095 GATGTAGAGACAAAGTGGCCCATCTGTCTACAGCAAACTTAACAAGATTCGCCAGG 9154
 QY 2768 ySerProHisLysThrGluThrTyrArgHisGlyProValIleProGluArgSerGluMe 2788
 Db 9155 GHTCTCT-----TTGCTAGTGCAGTGTTCAGAGCAATCTCCAT 9196
 QY 2788 tProSerLeu-----HisLeuGlySerProLysLysSe 2799
 Db 9197 CCTTCTGTACTGAAGAGGTTATCATCTGCGCAAAATTAAGCTTCAGCAAGAGCGCA 9256
 QY 2799 rGluSerSer-----ThrLysArgValValSerProAsnArgSerGluLeuLys 2815
 Db 9257 AAGATCCAGTGAATATGGAGAAATGCTAGAGGACCAACACCTGCTACCCAGAGAGCTT 9316
 QY 2815 rSer-----GlnLeuValMetSer-----ProGly-LysThrGlyMetH 2828
 Db 9317 TTCTAAAAAACCAAGAAAGCTCATGATGATGATTCACCTCGCAGAAAGACA----- 9369
 QY 2828 sLysHisIleLeuSerProSerLysValGlyLeuHisLys-----LysArgAlaLeuS 2846
 Db 9370 --CGGAAGGTACTGATTTGAGCCAGAGGAGACTCCAGAAAGTTGTAAGAAAGGCTTGC 9427
 QY 2846 eTrProAsnArg-----SerLysMetProThrGlnHisV 2857
 Db 9428 TGACATCCCGAGACAGAAAGACTAGCCCATATCTCTGCGAAGAACAMCATGTGGCAA----- 9483
 QY 2857 aIleSerProGlyLysThrGlyLeuHisLysAsn-----LeuThrG 2871
 Db 9484 --CTGGAGACACCCCGCTGCTGCACAGAAATGATAGCGGTATCCCGCATAGCTCGG 9541
 QY 2871 LuSerThrLeuPheAspAsnLeuSerSerProCysLysGlnIleLysValGlnIleAsnL 2891
 Db 9542 CAAGAAATTC-----TTGCAAGAGTCTCCAAACCAACAGCTGGTGGA-----G 9586
 QY 2891 euAsnSerProLysGlyLysLeuPheAsp-----ValLysSerL 2904
 Db 9587 CAGATCACAAAGATCAAAAGTCTCAGCGGAGCCAGTATGATTCAGGCACCATCTCTCG 9646
 QY 2904 ySerMetProTyrCysProSerGln-----PhePheAspAsnSerLysLeuGlyA 2921

```

Db 9647 AGAACCCACACGAAATCCCTCCAGTCAATTAATCTTCGAGAGAGTCCGACGTACAG 9706
QY 2921 sphaeseriguleuAsnThrLagLuserAsnAspLysSerGlnAlaGluAsnTrp 2939
Db 9707 CCCCAGAGAGGCGCTGAGGTGTCAGAGAGGCCACCTTGTCCAGCCCAAGCTGG 9762

RESULT 3
US-08-328-254-5
Sequence 5, Application US/08328254
Patent No. 5710022
GENERAL INFORMATION:
APPLICANT: Zhu, Xueliang
APPLICANT: Lee, Wen-Hwa
TITLE OF INVENTION: A No. 5710022el Nuclear Mitotic Phosphoprotein
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/328,254
FILING DATE: 24-OCT-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/141,239
FILING DATE: 22-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-CJ 1191
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 8789 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 544..7990
US-08-328-254-5

Alignment Scores:
Pred. No.: 7.6e-76 Length: 8789
Score: 1061.00 Matches: 625
Percent Similarity: 39.11% Conservative: 488
Best Local Similarity: 21.96% Mismatches: 936
Query Match: 7.18% Indels: 798
DB: 1 Gaps: 118

US-09-150-867-1 (1-2954) x US-08-328-254-5 (1-8789)
QY 529 LysGluAsnGlnLeuGlnTyrLeuProLysAspSerGlyAspMetAlaGluCysArgLys 548
Db 460 AAAAAGAAAGAAATATGAAGAAATG---AAAGAGAGAAAACTCTGTTTCTTGTGGAAA 516
QY 549 AlaserphegIuLysGluIleThrSerLeuGln-----GlnGlnLeuGlnSer--- 564
Db 517 AGTGAAGAAAGAAAACTTTTAACTCAGATGCAATCAGAAAGAAAACTTCCAGAGTAA 576
QY 565 -----LysGluGlnGluLysGln 571

```

```

Db 577 ATTAATCACTTGGAAACTTGCTCTGAAGACAGCAAAATATAAAGTCATGATACAGAG 636
QY 572 LeuValGlnSerPheGluLeuLysIleAlaGluLeuGlnLeuSerValLysAla 591
Db 637 AGAGTAAGAAACCTGAGAGATGAGAC-----AGAGAAACCTGAAGTCTCAGATC 684
QY 592 LysAsnLeuGluMetValThrAsnSerArgGluHisSerIleAsnAlaGluValGlnThr 611
Db 685 AGAAACCTTCACAAAGCTGTAGACAGTAG-----TAGGTGAGGTAGAGCCAGAAA 738
QY 612 AspValGluLysGluValAlaArgLysGluMetSerValLeuGluAspSerGlyTyr--- 630
Db 739 CTAGCTTATATGAGCTACAGCAAGAAA-----GCTGAGTCTCAGATCAGAAACATCAG 792
QY 631 ---AsnAlaSerAsnSerPheGlnAspSerSerValAspGlyLysArgLysSer 649
Db 793 AAGGAATAGAAATATGTTGTGAAGACTTCTCAGCTTACTGGGCA----- 840
QY 650 SerHisAspGluCysIleGluHisArg---LysMetLeuGluGlnLysIleValAspLeu 668
Db 841 -----GTTGAAGATCTAGACCAAGCTTCACTCTCAATGAATATATGACAAA 894
QY 669 GluGluPheIleGluAsnLeuAsnLysLysSerGluAsn-----AspLysGlnLysSer 686
Db 895 GACCGGTGTACCAACACTTGCATGCCAATATGAGAGCGTCAGGAGATCTGTAATAATCC 954
QY 687 SerGluGlnAspPheMetGluSerIleGlnLeuCysGluAlaIleMetAlaGluLysAla 706
Db 955 AAGATGCTTCTCTGTGTCAAATGAATGATCATCAGAGAAAGCTTTTG----- 1002
QY 707 AsnAlaLeuGluGluLeuAlaLeuMetArgAspAsnPheAspAsnIleLeuGlu--- 725
Db 1003 ---GTTTGTATCAGACGCTGCCATCATCTTCTTCAAAATATATGAGAACAA 1059
QY 725 ----- 725
Db 1060 GGAAGATGCTTCAGAGAGGAGTGAATGCTTTTGAAAGCAGCAAAAGTCCGAAAT 1119
QY 726 AsnGluThrLeuLysArgGluIleAlaAspLeuGluArgSerLeuLysGluAsnGlnGlu 745
Db 1120 TCTGCCATCTTACAAATATAGATGATTCATCTGAATTTTCAATAGCTCAAAAACAG 1179
QY 746 ThrAsnGluPheGluIleLeuGluLysGluThrGlnLysGluHisGluAlaGlnLeu 765
Db 1180 ATGAAC-----TCAGACTGCAAAAGCAGGTGA---GAGTTGGTG 1218
QY 766 HisGluIleGlySerLeuLysLysLeuValGluAsnAlaGluMetLysArgAsnGlnAsnLeu 785
Db 1219 CAATCAAGAGAGAAATATAGAGAAATCTCATGAAAGCAGAAACAGATGCAATCTTT 1278
QY 786 GluGluAspLeuGluThrLysThrLysLeuLeuLysGluGlnGluIleGlnLeuAlaGlu 805
Db 1279 GTG-----GCTGAACCAAGTCAAGCGCATTAATTAAGTTACAGAACACTTCTGTCAC 1332
QY 806 LeuArgLysArgLysAspAsn-----LeuGlnLysLysValArgAsnPheAspLeu 822
Db 1333 CAGAAATGTTCTGTGAACCTTAAGTCCCTTGAGAACAGAAAGAGCTGCAACTT 1392
QY 823 SerValSerMetGlyAspSerGluLysLeuGluGlnIlePheGlnLeuLysGlnSer 842
Db 1393 TTAAATGATATAGATAGAAATGAGCAG-----GCAAGATTCAGAAATTAATAAAGAGC 1446
QY 843 LeuSerAspAlaGluAlaValThrArgAspAlaGlnLysGlnLysSerPheLeuArgSer 862
Db 1447 -----AACCATCTACTTGAAGACTCTCTTAAGAGAGGTACAACTTTTA---TTC 1491
QY 863 GluAsnLeuGluLeuLysGluLysMetGlyAspThrSerAsnTrpLysArgGlnLysGlu 882
Db 1492 GAAACCTTAAGCTTGGAAGAAAGAAATAGATGTCATCTTCTCAATATAAAGGAA 1551
QY 883 LysAlaAlaSerLeuPheGluLysGlnLeuGluThrGluLysSerAsnTrpLysLysMet 902

```

Db 1552 -----ATTGAAGCTGACCCAGAGAAATGGACTTTAAGAAATT 1593
Qy 903 Glu1AspLeuGlnLysGlu-----LeuGlnSerAlaPheAsnGlnLysLeu 918
Db 1594 AATGATCCTTAAATCAAGAGAGATGAACTTAATCCAGAAAAGTAGAGTTTGCACAC 1653
Qy 919 TyrLeuAsnGlyLeuLeuAlaGlyLysValProArgAspLeuSerArgValGlnLeu 938
Db 1654 TATATAGAT-----GAAAG 1668
Qy 939 GluLysLysValSerGlnPheSerLysGlnLeuGln----- 950
Db 1669 GAGAAAAGCATTTGAGAGTTATCTGATCAGTACAGCAAGAAAACCTATTACTACAA 1728
Qy 951 -----LysAlaLeuGln 954
Db 1729 AGATGTGAAGAAACCGAATGCATATGAGATCTTACTCAAAAATACAAAGCACACAG 1788
Qy 955 GluLysAsnAla-----LeuGlnAsnGlnValThrCysLeuSerGlnTyrLys 970
Db 1789 GAAAAGATTTCTAAATTAGAATGCTTGTCTAAATGAAATGCACACTAGCTTTGGAATAAG 1848
Qy 971 PheLeuProAsnGlnValGluCysLeuLysAsnGlnIleSerLysAlaSerGlnLys 990
Db 1849 -----AAAAATGAGTTGGAAACAGCTAAAGACCATTTGCAAGGAACACCAAGAAATTC 1902
Qy 991 MetLeu-----LeuLysGlnGluGlyGlnHisSerAlaSerIleIleSerLysGlnLys 1009
Db 1903 TTAACAAATTAATGATTTGCTGAAAGAAAGAAATCAGAAATGCTGAGTTGAGACA 1962
Qy 1010 IleMetGln-----GluGlnSerLys 1016
Db 1963 GTGAGAGACAGCTGAGATCTGAGATGACATACCAAAACAAATTCAGAGCGACGCT 2022
Qy 1017 -----GlnIleLeuGlnLeuThrAspGlnValThrHisThrGlnSerLys 1031
Db 2023 GGTGGTTTAAAGCAAGAAATCATGACTTAAAGCAAGAAACAAATTCAGAAAGCA 2082
Qy 1032 Val-----GlnGlnThrGlnGlnGlnIleTyrIleGlnMetLysLysMetHisAsp 1047
Db 2083 GTTAATGACTTATTACAAAGAAATGAAACAGCTGATGAAGGTAAATGAACTTAACATGAA 2142
Qy 1048 -----AspLeuPheGlnLysLysThrLysAsnLysSerGlnLysAlaGlnAspLeuArg 1065
Db 2143 TGTCAAAATCTAGAAATCGAAACCAATTAAGAACTCTGAAAGAAAGAGAGAGTACAGA 2202
Qy 1066 GluMetGlnAsnLeuLysGlyThrMetGln----- 1075
Db 2203 AATCAATGTATTTTAAACCTCAGATGATCTTGAAGTTAAAGAAATTTCTAGATAGT 2262
Qy 1076 -----SerValGlnValLysIleAlaSerPheLysGlnLeuGln 1089
Db 2263 TATAATGCGCAGTTGTGCAATTAAGAACTATGCTAAGAAATTAAGAAATTTAAAGTTAG 2322
Qy 1090 GluThrIleArgAspLysGlnGlnLeuLeuHisGlnLysLysTyrPhePheGlnAlaMet 1109
Db 2323 GAAAGTGAAAGAGAGAGAGATGCTGCATGAA-----TTA 2361
Qy 1110 GlnThrIlePheProIleThrProLeuSerAspSerLeuProProSerLysLeuValGln 1129
Db 2362 CAGCAATTT-----AGAGAGATCTTGAACCCAGCAATTTGCANAGC 2403
Qy 1130 GlnAsnSerGlnAspProIleGlnIleAsnAspTyr----- 1141
Db 2404 ATGCAAGTCACAAAGAAATTAAGTGGCCTTAAGACATGTCAAATAGATCCGGAAGAAAGTAT 2463
Qy 1142 -----HisAsnLeuIleAlaLeuAlaThrGlnLysAsnAsnIleMetValCys 1157
Db 2464 ATTTCAGGCGCCTCATGATGTGTCAACAAGTCAAAACGACATGCACTTCAGTGCCT 2523
Qy 1158 LeuGlnLysGlnLysAsnSerLeuLysGlnValIleAspLeuAsnThrGlnLeuGln 1177
Db 2524 CTGCAAAACAAATGAACAGCTGATGAG-----CTAGAGAAATAATGTGAA 2571

Qy 1178 SerLeuGlnAlaGlnSerIleGlu-----LysSerAspLeu 1189
Db 2572 ATACTGACAGGCTGAAAGATGATAGCTCTTAAGTACAGTGAATGATTCAGAGTCAAGATGT 2631
Qy 1190 GlnLysProLysGlnAspLeuGlnGlnGlnValLysLeuLeuGlnLysMetGlnLeu 1209
Db 2632 ATCAGACGACACGAGAAATATGCGACAGAGAGTAGGGAACCTCTAATGAACTTAAATTA 2691
Qy 1210 LeuLysGly----- 1212
Db 2692 TTAATGATGACAGTGTCTTCTCCATGCTGATGATTAGTGAAGACATACAGAGGTGAA 2751
Qy 1212 ----- 1212
Db 2752 TTGGTGAACCAACCAATGAACAGCACCTGTCTTTGGCTCCATTTGAGCAGAGTAAAT 2811
Qy 1213 -----HisLeuThrAspSerGlnLeuSerIleGlnLysLeuGlnLeuGlnLysLeu 1229
Db 2812 TCTTACGAGCACTTGACA-----TTGTACAGCAAGAAAGTTCAAATGCACCTTGGC 2862
Qy 1230 GluValThrGlnLysLeuGlnThrLeuGlnGlnGlnMetLysAsnIleThrIleGlnArg 1249
Db 2863 GAATTCAGAGAGCAAAATCTTATCTTACAAAGTGAACCAAAATTTTACATGATCAGCAC 2922
Qy 1250 AsnGlnLeuGlnThrAsnPheGlnAspLeuLysAlaGlnHisAspSerLeuLysGlnAsp 1269
Db 2923 TGTCAATGAGCTCTTAATATGACAGAGCTGCAGACCTATGTTGACTCATTAAG----- 2976
Qy 1270 LeuSerGlnAsnIleGlnGlnSerIleGlnThrGlnAspGlnLeuArgAlaLeuGlnGln 1289
Db 2977 ---GCCGAAATTTGGTGTCTGCA-----ACGAATCTGAGAACTTCAGAGT 3021
Qy 1290 GluLeuArgGlnGlnLysGlnLeuValAspSerPheArgGlnGlnLeuLeuAspCysSer 1309
Db 3022 GACTTGTGGAAGAGATCAGCTG-----GGCTTGGAAGAGGGCTGCTTCATCCCTG 3075
Qy 1310 ValGlyIleSerSerProAsnHisAspAlaValAlaAsnGlnLysValSerLeuGly 1329
Db 3076 TCATCTCTTGTGTGCTGCACAGCTCTAGCTTAC-----AGTTTGGGA 3120
Qy 1330 GluValAsnSerLeuGlnSerGlnMetLeuArgGlnLysArgAspGlnLeuGlnThrSer 1349
Db 3121 GAC---TCTCTCTTTTACAGAGCTCTTTAGAA-----CAGACAGGA 3159
Qy 1350 CysLysAlaLeuValSerGlnLeuGln-----LeuLeuArgAlaHisValLysSerValGln 1368
Db 3160 GATATGCTCTTTTGAATTTAGAGAGGGCTGTTTCAGCAAAACAGTGCAGTATAGT 3219
Qy 1369 GlnGlnAsnLeuGlnIleThrLysLysLeuAsnGlnLysGlnLysGlnIleLeuGlnLys 1388
Db 3220 -----GAAGTATTT-----TGCACAGCTGTGCAGAGCAAAATTTGACCCAGG 3261
Qy 1389 SerGlnGlnSer-----GluValIleLysSerMetLeuGln 1401
Db 3262 AAAGAAACCCCTTGGCCCCAGGAGGAGGTGTGAAGAGCTTGAATCCCTCTGTAGAGTG 3321
Qy 1402 LeuLysGlnAspAsnAsnLysLeuLysGlnGlnAlaGlnGlnLysSerSerLysGlnAsn 1421
Db 3322 TACCGGAGCTCCCTGAGAACCTGAAGAGAAATGAAGATCAAGGATTAATGAAATAT 3381
Qy 1422 GlnPheSerLeuGlnGlnValPheSerGlnLysLeuValAspGlnIleGlnVal 1441
Db 3382 AAG-----GAAATTCAGAG 3396
Qy 1442 LeuLysAlaGlnLeuLysAlaAlaGlnLysGlnGlnLysAspArgAspTyrPhe 1461
Db 3397 CTCGAGCAGTATTATTAAGTTCTGAAGAGCAGCTTACCTGATGAGACAGATATTG 3456
Qy 1462 GluLeuValGlnThrAlaAsnThrAsnLeuValGlnGlyLysLeuGlnThrProLeuGln 1481
Db 3457 TCA-----GAA 3462

QY 1482 AAlaSPHISgluLnuSPserIleasparGargSerGluGluMetGluIleLysValLeu 1501
 DB 3463 AATGAACAGTGGCAACAGAGAGCTGACAAAGCTGTGAGATGAGTGCACAGTGGCG 3522
 QY 1502 GlyGluLysLeuGluLnuArgSngInIlyrLeuGlu-----ArgLeuGlu 1517
 DB 3523 GCAGAAAGAAACAGACGAGACACTGTGACTTGAGCTGGAAGTAGACAGATCCAGCA 3582
 QY 1518 GluLysLeuGluLeuSerSngInLysLeuGluIleLysGluMetGluThrSerVal 1537
 DB 3583 CAAGGTCTGAGCTTAAGTCTCTCG-----TCT 3609
 QY 1538 LeuLeuLysAspSpleuGlnGluLysLeuGluSerLeuLeuSerGluSngInIleLeu 1557
 DB 3610 TTGCTGGCAGTCACAGAGAGATGCTATTCAAGGC-----CGA 3648
 QY 1558 LysGluSngInLysPThrThrLeuLysSngInLysSerAspThrGluAlaGlnLeuGluLys 1577
 DB 3649 AATGAGAGCTGTGACATATCAAAAGAAACATATTCTCAGAAACTCAGAAAGAACCCAAAG 3708
 QY 1578 -----ThrGlnGluLysLeu----- 1584
 DB 3709 CATGATGTCATCAGATTTGTGATAAAGATGCTCAGCAGAGCTCAATCTAGACATTGAG 3768
 QY 1585 -----AlaLysSngInLysAlaIleAlaSerAspSngInLysProIleThrGlnGluLys 1602
 DB 3769 AAAATAACAGAGCTGTGCTGATGAACCCACAGAGAGTGTCTGGGAGAACGTCCCA 3828
 QY 1603 GluThr-----SerAlaAspCysValHis 1610
 DB 3829 GATACCAATTATGAGCTCCAGGGAGAGATAAACCCAGGCTCTCAGATGATCATTTCT 3888
 QY 1611 ProLeu-----GluGlu 1614
 DB 3889 GAATTGCATTTCTGCTCTCTAATGCTTTGCTGATGCTGATTTCTCGGGAGATCAGGA 3948
 QY 1615 LysIleLeuLeuLeuThrGlnGluLysLeuGlnLysThrSngInGluLysLeu 1634
 DB 3949 GATATCATATATCTTCAACTGCGGGTAAAGACATCAATGAGATTTGAGATTCTT 4008
 QY 1635 HisGluLysSngInLysLeuGlnAlaGlnValGluLeuLysCysGluValGlnHisLeu 1654
 DB 4009 CATGTGATAGAGACCTGACAG-----AAAGTTGAAGTTG 4047
 QY 1655 MetLysSerMetIleGluSerLysSerSerLeuGluSerLeuGlnHisGluLysHisAsp 1674
 DB 4048 CTAAATGAATGAAGAATGAGTCAAACTC-----CATTTA 4086
 QY 1675 ThrGluGlnGluLeuAla-----LeuLysGlnGlnMetGln 1687
 DB 4087 CAGGAGGTACAACTAAATGATGCCAAATGAAACATGATGAAATGAAATGATGGG 4146
 QY 1688 ValValThrGlnGluLysGluLeuGlnInThrHisGlnHisLeuThrAlaGluVal 1707
 DB 4147 GAACTTAAGAAAGAAACTCAGATTTAAGTAAATGGAATTTTCTTGAGATCAG 4206
 QY 1708 AspHisLeuLysGluAsnIleGluLeu-----GlyLeuAsnPheLysAsnGlnAlaGln 1725
 DB 4207 CAGAGGTACTCCAGAGAGTAGAAACTTCTGAGGCTCAATTCGATTGAAATGAT 4266
 QY 1726 ---GlnLysThrThrLysGluGlnCysLeuLysSngInLys----- 1739
 DB 4267 GCAGATTAATCATCAGCTGAGATATTGAGATATGTGGCCAGAGTGAATGACACTGG 4326
 QY 1740 -----GluLeuGlnGlnSerGlnHisArgLeuGlnCysGlu----- 1751
 DB 4327 AAGGAGATTTCTTGATGTGAAAAATGAGCTGAGTACAGATCAGATCGGAGAAAGCTAGC 4386
 QY 1752 -----IleGluLysMetLysSerLeuLysAspSngInSer 1764
 DB 4387 AATGAGCATGAAGCCCTGACTGAGGCTGACTTAAAGAGTACTTCAAAACAGAGAACTA 4446
 QY 1765 AlaLeuGluThrLeuLysGluSerGluLysValIle---AsnLeuAsnGlnGlnMet 1783

DB 4447 TGTATTAGAAAAGACAAATGAATAAGCAGAGGTATTGTGCTGCTGAGAGAACTC 4506
 QY 1784 GluMetValMetLeuGluMetGluLysSngInSerGlnArgThrAlaIleAlaGlu 1803
 DB 4507 TCAGTGGTCACAAAGTGAAGAAACAGCTTGTGGAGATTAAGTACTATGTCAAAAAA 4566
 QY 1804 Arg-----AspGlnLeuGlnAspAspLeuArgLysSerValGluMetSerIleGlu 1820
 DB 4567 ACCACGGACACGATTCATCTGTGAAAAAATGAAGGAAACACAA---GAGCTTGAG 4623
 QY 1821 Thr-----GlnAspAspLeuArgLysAla 1828
 DB 4624 TCTCATCAAGTGAAGTGTCTCCATTGATTCAGGTGGCAGGCGAGGTAAAGAAAG 4683
 QY 1829 GlnGluAlaLeuGlnGlnGlnLysAspLysValGlnGluLeuThrSerGlnIleSerVal 1848
 DB 4684 ACCGAACTCTTCAGACTTGTCCCTGATGATGAGTGAAGTAAAGACAAACATCAT 4743
 QY 1849 LeuGlnGluLysIleSerLeuGluLysSngInMetLeuTyraSngInAlaThrValLys 1868
 DB 4744 CTCGAGAAAGAGCTCAGAGTTGGAAAAAGACCTCACAGGACATGTCTTTGACAAATGT 4803
 QY 1869 GluThrLeuSerGluArgAspSpleuSngInSerLysGlnHisLeuPheSerGluIle 1888
 DB 4804 GAGCTGAAAACCAATTCACACACTGATGATTAAGAAAGAAATTCCTTCTCAAGGAATCT 4863
 QY 1889 GluThrLeuSerLeuSerLeuLysGluLysGluPhe-----Ala 1901
 DB 4864 GAAAGCTCCAGCCAGCACTGATGATCAGATTTGAAAAAGCTGATGCTCCAGGCC 4923
 QY 1902 LeuGluGlnAlaGluLysAspLysAlaSerAlaAlaArgLysThrIleAspIleThrGlu 1921
 DB 4924 TTGAGGCCGACGTGTGGAGAAAGTGTGCGATTGAGCTGAGCTCAACACAGAG 4983
 QY 1922 LysIleSerAsnIle-----Glu 1927
 DB 4984 GAAGTGCATCAGCTGAGAGAGCATCGAAGAACTCAGAGTTCGCAATGAGGCCATGAA 5043
 QY 1928 GluGlnLeuGlnGlnAlaThrAsnLeuLysGluThrLeuTyraArgLysSerLeu 1947
 DB 5044 AAGAACACCTGCACATTCACAGAAACTGAAGAACCGCAGCGGAGATATTCACCTT 5103
 QY 1948 IleGlnCysLeuGlnLeuAlaLeuAsnThrGlnHisIleLeuArgGluThrLeuLysSer 1967
 DB 5104 -----AAGATTAAGTTGAGACCTGAAAGGAAATTCAGATGTCAGAAACAAAC 5154
 QY 1968 LysAspLeuAlaLeuGlnLysMetGlnGlnTyraArgSpleuAlaAlaSngInValIle 1987
 DB 5155 CAGGAGCTAGTGAT---CTTGATGCCAGAAATTCCAAAGCA----- 5193
 QY 1988 AlaLeuThrGluLysMetSerSerLeuGlnGlnIleAsnGlnLysValThrThrLeu 2007
 DB 5194 -----GAGTGAAGACTCTAAACAACTGAAGACATGGCCAGAGCCCTG 5241
 QY 2008 LysGluGluGlu-----GlyLysGlnThrPheTyraLeuGln 2020
 DB 5242 AAAGTTTGAATTAAGACCTTGCAGGTTAAGGTCTGAAAGAAATATGACAAACAA 5301
 QY 2021 ArgProSerLysGlnGlnSerSerSerGlnMetGluGluLeuArgLysSerLeuLysThr 2040
 DB 5302 ATCAACAAACAAAGCTCAGTTGTCAGAACTGAGAACAGTTACTCTCTCATTTAAAG 5361
 QY 2041 LysAspLeuGlnGlnGlnAlaGlnLysGluIleSerGluAlaThr----- 2056
 DB 5362 CTGTTAGAAAGAAAGGAGCAGAGATACAGATCAAGAAACAAATCTAAACTGCAAGTC 5421
 QY 2057 -----AsnGlnIleLysAsnLeuThrAlaLysIleSerSerLeu-----Glu 2070
 DB 5422 GAGATGCTTCAGATATCAGTTAAAGAGACTTAAAGAGCACTAGACAGCTGTGTGTGATC 5481
 QY 2071 GluGluIleLeuGln---AsnAlaSerIleLeuSngInLysValAlaSerGluValGluAsn 2089

Db 5482 CAAGAAATATGAAAGCCAGAACAGAGCTAGACCCAGCAATAGAGAAAGAGCATG 5541
 QY LeuArgHisSerLysGlnGlnIleuValSerGluLeuGlu-----GlnLeu 2104
 Db 5542 CTGAGAAATAGCATTTGAAAGCTGAGAGCCCGCTAGAGCTGATGAAAGAGAGAGCTC 5601
 QY 2105 SerLeuThrLeuLysSerArgAspHisAlaPheAlaGlnSerLysArgGluLysAspGlu 2124
 Db 5602 TGTGCTTTACAAACAACGTGAG-----GAAAGTGACCATCATGACAGATTTA 5646
 QY 2125 AlaValAsnLysIleAlaSerLeuAlaGluIleLysIleLeuThrLysGluMetAsp 2144
 Db 5647 CTTAAGGATGAGTGGAGAACTTGAAAGAGAGCTAGAGATAGCCAGAGCAAAACCAAG 5706
 QY 2145 -----GluPheArgAspSerLys-----GluSerLeuGlnGluGlnSer 2157
 Db 5707 CATCAGCTCTTGAGGCGAGACAAATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5766
 QY 2158 SerHisSerSerGluGluLeuGluLeuGluLeuGluLeuGluLeuGluLeuGluGln 2177
 Db 5767 GAAGGATGAGCCAAAGCTGAGAGGCTGAGATTAAGATGTTGTTACTATTAAGTCAGAA 5826
 QY 2178 LysGluAspIleAsnAsnLysLeuAlaGluLysValLysGluValAspGluLeuGln 2197
 Db 5827 AAAAATAATCTGACAAATGAAATTCAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5880
 QY 2198 HisLeuSerSerLeuLysGlnLeuAspGlnIleGlnMetGluLeuArgAspGluLys 2217
 Db 5881 -----GlnPheArgAspSerLys-----GlnPheArgAspSerLys 2237
 QY 2218 LeuArgAsnArgGluLeuGluLysMetAspIleMetGluLysGluIleSerValLeu 2237
 Db 5899 TTTGAAAT-----ATTG 5913
 QY 2238 ArgLeuMetGlnAsnGluProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 2257
 Db 5914 CAAGAAAG 5973
 QY 2258 LeuGluSerArgAsnGlnIleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 2277
 Db 5974 CTTCAGAA-----CAATTAAAGAGCTCATGAGAGAGAGAGAGAGAGAGAGAGAG 6024
 QY 2278 GlnGlnHisThrLeuLeuSerSerLeuSerSerGluLeuGlnLysGluThrGluAlaHis 2297
 Db 6025 GACCA-----GAGCCCTG 6039
 QY 2298 LysHisLysMetLeuAsnIleLysGluSerLeuSerSerThrLeuSerArgSerPheGly 2317
 Db 6040 AAG-----GCCAAAG 6075
 QY 2318 SerLeuGlnThrGlnHisValLysLeuAsnThrGlnLeuGlnThrLeuLeuAsnLysPhe 2337
 Db 6076 TGTCTTAACCTGAG 6135
 QY 2338 LysValValIlyrArgThrAlaAlaValLysGluAspHisSerLysIleLysAspTyrGlu 2357
 Db 6136 ATTTGTTTGCATCTCAGAG-----AATGCTCATTCAGAGAGAGAGAGAGAGAGAG 6180
 QY 2358 LysAspLeuAlaAlaGluGlnLysArgHisAspGlu-----LeuArgLeuGlnLeu 2374
 Db 6181 GATGCAAG 6240
 QY 2375 GlnCysLeuGlnGlnHisGlyArgLysTrpSerAspSerAlaSerGlnGluLeuLysPhe 2394
 Db 6241 CAAGAGCAAG 6300
 QY 2395 -----CysGluIleGluPheLeuAsnGluLeuLeuPheLysValAsnIle 2410
 Db 6301 AAG 6360
 QY 2411 IleIleSerValGlnAspArgPheSerGluValGlnValPheLeuAsnGlnValGlySer 2430
 Db 6361 CTACAAATCAAAAAGGCTCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6420

QY 2431 ThrLeuGlnGluLeuLeu-----HisLysLysGlyPheMetGlnTrpLeuGlnIlePhe 2449
 Db 6421 AATCTAGAGATGACCTTGAATTGACAAAATGACAAAATGCTCTTGTGAAAAGCTA 6480
 QY 2450 Gly-----AspLeuHisValAspAlaLysLys 2458
 Db 6481 AACAAATGACTGCAAGAAAG 6540
 QY 2459 -----LeuSerGluGlyMetGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 2475
 Db 6541 ACAG 6600
 QY 2476 LeuLeuThrLysArgLeuLysAlaValAlaGlnSerLysIleGlnArgGluIleThrVal 2495
 Db 6601 TTACTGTTGGAAGAAATTAAGAG-----ACCAAGATCA----- 6636
 QY 2496 TyrLeuAsnGlnPheGluAlaLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 2515
 Db 6637 -----TTGAAG 6672
 QY 2516 ArgArgMetGlnHisGlyProSerAlaSerValMetGluGlnGlnGlnGlnGlnGln 2535
 Db 6673 AAG-----AGCTAG 6711
 QY 2536 LeuGlyIleLeuLysThrValGlnAspGluSerLysLysLeuGlnSerArgLysMet 2555
 Db 6712 GAAGG-----AAAGTGAG 6753
 QY 2556 LeuGluAsnGluLeuAsnLeuValLysAspAlaMetHisLysGlyLys----- 2573
 Db 6754 -----CTTCAG 6777
 QY 2574 ValAlaIleLeuGlnAspLysLeuLeuSerArgAsnAlaGlnAlaGluLeuAsnAlaMet 2593
 Db 6778 CAGGCTTTGCTTTGAG-----ACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6831
 QY 2594 GlnValLysLeuThrLysLysGlnAspAsnLeuGlnAlaAlaMetLysGlnIleGluAsn 2613
 Db 6832 CGAGAGAAATGACTTCTTAAG 6891
 QY 2614 LeuGlnLysMetValAlaLysGlyAlaValProTyrLysGlnGluIleAspAsn----- 2631
 Db 6892 TTAAGCTAG-----AAAG 6927
 QY 2632 -----LeuLysThrLysValValLysIleGlnMetGlnLysIleLysTyr 2646
 Db 6928 AAGCTACTACTAGATTGGAAGAAATGGAAGAAACCAAGAGAGAGAGAGAGAGAGAG 6987
 QY 2647 -----SerLysAlaThrAspGlnIleAlaLysLys 2658
 Db 6988 GTAATACGTTGAG 7047
 QY 2659 SerCys-----LeuGluAspLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 2676
 Db 7048 TCCGTGTAACAGCTGGAAG 7107
 QY 2677 ArgAlaGlnAlaAspAsnAspThrThrValLysValProLysAspTyrGlnLysAlaSer 2696
 Db 7108 GCTGACAG 7131
 QY 2697 ThrPheProValThrCysGlyGlySerGlyIleValGlnSerThrAlaMetLeuVal 2716
 Db 7132 -----GTTACTGTTATGATACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7161
 QY 2717 LeuGlnSerGlu-----LysAlaAlaLeuGlu-----ArgGluLeuSer 2729
 Db 7162 TTAACACCTGAGATCAAG 7221
 QY 2730 HisTyrLysLysGlyTyrHisHisLeuSerArgThrMetSerSerSerGluAspArgLys 2749
 Db 7222 GAATACCTTGATAGTACTGTTCTTG-----CTTAATAGAGAGAGAGAGAGAGAGAGAG 7272

QY 301 AAlaYThrValIleIleCysThrIleThrProValSer-----PheAspGluThrLeu 318
 DB 994 GCCAAAGACCCATATATGCAAAATGTTTCATCTCGATGCTGGTGGGGAACCCCTTA 1053
 QY 319 SerThrLeuGlnPheAlaSerThrAlaYHisValArgAsnThrProHisValAsnGlu 338
 DB 1054 TCACACTACTCTTGGCTCAAGAGCCCAAGCTGATTAATAAACAGGCACTGATAATGAA 1113
 QY 339 ValLeuAspArgGluAlaLeuLeuLeuYArgTyrArgGluGluIleLeuAspLeuYAs 358
 DB 1114 -----GACACCCCAAGAAATATGAGCCCAAGCTCCAAAGTGAAGTGAAGCTCCAAAGAA 1167
 QY 359 GlnLeuGlnAsnLeuGlnLeuSerSerSerGluThrIleGlnValAlaMetAlaGluGlnGlu 378
 DB 1168 CAACTGGCGAGCTTGGCTTCAGACAGACACCCCAAGAACTCTCCACAGAGACAAA 1227
 QY 379 HisThrGlnLeuLeuAlaGluIleLeuGlnLeuHisYAspGluAspArgIleTyr 398
 DB 1228 AAGAAGACTACTATATGAGCTATTTCCAG-----GAAGCAATGTTATTC 1272
 QY 399 HisLeuThrAsnIleValAlaAlaSerSerGlnGlnLeuSerGlnGlnAspGlnArgValYAs 418
 DB 1273 TTT-----AAGAAATCTGAACAGGAAGAAAG----- 1299
 QY 419 ArgYAspArgValIleThrPAlaProGluYAspIleGlnAsnSerLeuHisAlaSerGlu 438
 DB 1299 ----- 1299
 QY 439 ValSerAspPheAspMetLeuSerArgYLeuProGluYAsnPheSerYAspAlaYAspHe 458
 DB 1300 ---TCTCTGATAGAAAAGATTACCAATTAGAACCTCACCCTCAAAAGAAAGAAATTT 1356
 QY 459 SerAspMetProSerPheProGluIleAspAspSerValCysThrGlnPheSerAspHe 478
 DB 1356 ----- 1356
 QY 479 AspAspAlaLeuSerMetMetAspSerAsnGluYleAspAlaGluTyrAsnLeuAlaSer 498
 DB 1357 -----ATTCATCTATATATAATGATTGCGAATTTCCGA---GAGCAT 1395
 QY 499 LysValIleThrHisArgGluYAspYThrSerLeuHisGlnSerMetIleAspPheGluGlnIle 518
 DB 1396 CAAATATATACCTTGGAAAG-----CTCCAC----- 1422
 QY 519 SerAspSerValGlnPheHisAspSerSerYAspGluAsnGlnLeuGlnTyrLeuProYAs 538
 DB 1423 -----AAGAAATCCCGGAGAGTTTCTGCTGAG 1452
 QY 539 AspSerGluAspMetAlaGluCysArgYAspAlaSerPheGlnGluYAspGluIleThrSerLeu 558
 DB 1453 GAGCAGATGCTTGCCTC-----TCAGAAATTAAGAGATGAGATTCAAACTCTG 1500
 QY 559 GlnGlnGlnLeuGlnSerYAspGluGlnGluYAspGluLeuValGlnSerPheGlnLeu 578
 DB 1501 CGAACAATAATAGAGCCCAAGCTCCAGATGCAAAAGTATGCTAATGAAAAATCATTTCCCTC 1560
 QY 579 LysIleAlaGluLeuGlnGlnLeuSerValLysAlaYAsnLeuGlnMetValIleThr 598
 DB 1561 AGG-----GAGGAGAAATAGAAAGACTGAGATTA-----TTAGACCTCTGTGAAA 1602
 QY 599 AsnSerArgGluHisSerIleAsnAlaGluValGlnThrAspValGluYAspGluValAl 618
 DB 1603 AGAGCTCAAGAA----- 1614
 QY 619 ArgYAspGluMetSerValLeuGluYAspSerGluYAsnAlaSerAsnSerAspLeuGln 638
 DB 1614 ----- 1614
 QY 639 AspSerSerValAspGluYAspArgYLeuSerSerSerHisAspGluCysIleGlnHisArg 658
 DB 1614 ----- 1614

QY 659 LysMetLeuGlnGlnYAspIleValAspLeuGlnGluPheIleGluAsnLeuAsnYAsp 678
 DB 1615 ---ATGATGCCACAGACCATTCGAAACAACTAGAAAAAGCTTTCTCGAATATAGTGGCATG 1671
 QY 679 SerGluAsnAspYAspGlnYAspSerSerGlnGlnAspPheMetGluSerIleGlnLeuYAs 698
 DB 1672 GAGAAATAATGACAAA-----AATCAGCAAGATTTTCCACCTAAAGCTCAGAAAGAG 1722
 QY 699 GluAlaIleMetAlaGlnYAspAlaAsnAlaLeuGlnGluLeuAlaLeuMetArgAspAsn 718
 DB 1723 CCATGTTGTTTGC----- 1737
 QY 719 PheAspAsnIleIleLeuGlnAsnGluAsnGluThrLeuYAspArgGluIleAlaAspLeuGluY 738
 DB 1738 -----AACACTGAAAGTTAAAGACAACTCTCCAAATTCAGCA 1779
 QY 739 SerLeuYAspGluAsn---GlnGluThrAsnGluPheGluIleLeuGluYAspGluThrGln 757
 DB 1780 GAGCTGAATTAATTCAAACAGAAATATGAAATTC-----AAGAACTTACT 1827
 QY 758 LysGluHisGluAlaGlnLeuIleHisGluIleGluYSerLeuYAspYAspGluValGluAsn 777
 DB 1828 AGGAAAGGCAAGCTAGAAATTCAGATCAGAGCTTCAGCTTTGCAAAAAGCG----- 1878
 QY 778 AlaGluMetYAsnGlnAsnLeuGlnGluAspLeuGluThrIleYAspThrYAspLeuYAs 797
 DB 1879 -----AACCTTAATCTTCAAAAACCTTTGGAAAGCA---ACAAAAGCTCGCAAG 1923
 QY 798 GlnGlnGlnIleGlnLeuAlaGlnLeuYArgYAsp---ArgAlaAspAsnLeuGln----- 814
 DB 1924 CGGCAGAAAGTT-----TTCAGCTGAATTAATTCGTCGAAACACTTAAGATTATA 1977
 QY 815 -----LysYAspValArgAspHe-----AspLeuSerValSer 825
 DB 1978 ACTACACCAACCAAGGCTTACCACCTTCAATCCGACCAAGTACCAAAATTAAGCCCTTGAA 2037
 QY 826 MetGluAspSerGluYAspLeuGlnGlnGluIlePheGlnLeuYAspGlnSerLeuSerAsp 845
 DB 2038 ATGGAGAACTTGGCTGCTCTA-----TACACTCAGAAATTCAGCATATTAGAT 2085
 QY 846 AlaGluAlaValIleThrArgAspAlaGlnYAspGluCysSerPheLeuYAspSerGluAsnLeu 865
 DB 2086 AATGATATATTAATAGACCCAGTCTCTCTGAG----- 2118
 QY 866 GluLeuYAspGluYAspMetGluAspPheSerAsnTyrArgGlnYAspGluYAspAlaAla 885
 DB 2119 -----ATGATGAACAAAGCTTT 2136
 QY 886 SerLeuPheGluYAspGlnLeuGluThrGluYAspSerAsnTyrYAspMetGluAlaAsp 905
 DB 2137 GAGGCCATTTCTGAAGACTTGAACAGTGAACAGTGAACAAAGAGTGCCTTCAAGCCAAA 2196
 QY 906 LeuGlnYAspGlu-----LeuGlnSerAlaPheAsnGluIleAsnTyr 919
 DB 2197 CTGATGTAAGAAAGCATTAATAACCTTAAGCTTACAGCATGTTGCAAAACTGCAATAT 2256
 QY 920 LeuAsnGluLeuLeuAlaGluYAspValProArgPheLeuSerArgValGluLeuGln 939
 DB 2257 CATTTACCCAAATG-----CAGAGCTTTCTCTATCTCAAGAAAGATTT--- 2298
 QY 940 LysYAspValSerGluPheSerGlnLeuGlnYAspAlaLeuGlnGluYAspAlaLeu 959
 DB 2299 -----GATTGACCAAAACAGAGAGAGCTTCTCAGACTGAATGATGCTT 2346
 QY 960 GluAsnGluVal---ThrCysLeuSerGluTyrYAspPheLeuProAsnGluValGluCys 978
 DB 2347 GAAAGCAGCTTCAAGACACTCAAACTAAATGACTTTTGAAGAGGAGCTACATGAC 2406
 QY 979 LeuYAsnGlnIleSerYAspGlnGluIleMetLeuLeuYAspGlnGlu----- 996
 DB 2407 CTGGGAGTAGTCCCTTCATTTGCTGACAAAGAGCTTCTTCAAGTAATTTGGAATTAAGT 2466
 QY 997 -----GluYAsnIleSerAlaSerIleIleSerYAspGlnGluIleMet 1011

Db 2467 TCATCAAAAGCAATCAGAGAAAGATTCAACAAACTTCCGAAAGACATCGATGTA 2526
 QY 1012 GlnGlnGlnSerGlnGlnGlnLeuGlnLeuThrAspGluValThrHisThrGlnSerLys 1031
 Db 2527 CACCTTCATTAATGATATATCTCAGGTAGAAAAGAAACCTCTGACGACAAACCTGTC 2586
 QY 1032 ValGlnGlnThrGlnGlnGlnGlnLeuGlnMetLysLysMetHisAspAspLeuGln 1051
 Db 2587 CTACAGAGATTCTATGACAACCTTACAGAAATATGAAATTTGACATTGACCACTTCA 2646
 QY 1052 LysTyrIleArgAsn---LysSerGluAlaGluAspLeuArgGlnMetGlnAsnLeu 1070
 Db 2647 AGAAACCTCCAAACCTTCAAAAAGAAATGAACCTGAAATCTGATCTGATTAATTG 2706
 QY 1071 LysGlyThrMetGlnSerValGlnValLys-----IleAla 1082
 Db 2707 ATGAGACTCTTGAGCGACAAAAGACGCAATTAACAATTAATCATACAGTTTGAAAGAA 2766
 QY 1083 AspThrLysHisGlnLeuGlnGlnThrIleArgAspLysGlnGlnLeuHisGlnLys 1102
 Db 2767 GATTAAGAAAGACAGTTCTTAAGAAATCTTAAGATTTTGAGGGCTGACGACGAGAA 2826
 QY 1103 ---LysTyrPhePheGlnAlaMetGlnThrIlePheProIleThrProLeuSerAspSer 1121
 Db 2827 CAGAAAGAGACGGCCAGCTGTGACGACGACGACGACGACGACGACGACGACGACGAC 2886
 QY 1122 LeuProSerLysLeuValGlnLysAsnSerGlnAspProIleGluIleAsnAspTyr 1141
 Db 2887 TTG----- 2889
 QY 1142 HisAsnLeuIleAlaLeuAlaThrGlnLysArgAsnAsnIleMetValCysLeuGlnThrGlu 1161
 Db 2890 -----CTGCTACTGCAAAAAA-----GTGATCAGCTGCTCCGTAAGAAAGTCT 2928
 QY 1162 ArgAsnSerLysGlnGlnValIleAspLeuAsnThrGlnGlnGlnSerLeuGlnAla 1181
 Db 2929 AGAGATTCTGATTAAGAAAGTGTAGCTGACCTCATGAACCGATCCAGAGACTTAAGTCA 2988
 QY 1182 GlnSerIleGlnLysSerAsp---LeuGlnLysProLysGlnAspLeuGlnGlnGlu 1200
 Db 2989 TCGGTCTGTGAGAAACGAAACATATAGACACCTGAAACAGAACGACATGAAT 3048
 QY 1201 ValLys-----LeuLeuGlnMetGlnLeuLeuLysGlnHisLeuThrAspSer 1217
 Db 3049 TGCATAATACACTCTGCTTGTGTTGACAGAA-----GAGAGC 3087
 QY 1218 GlnLeuSerIleGlnLysLeuGlnLeuGlnAsnLeuValThrGlnLysLeuGlnThr 1237
 Db 3088 AGAGTGTGATCAGAAAGAGAGAGATGATTTCTGATCTGAAGAAACCTTAGG--- 3144
 QY 1238 LeuGlnGlnGlnMetLysAsnIleThrIleGlnArgAsnGlnLeuGlnThrAsnPheGln 1257
 Db 3145 -----CTGAGAAATACTTCTCTGAG 3162
 QY 1258 AspLeuLysAlaGlnHisAspSerLysGlnAspLeuSerGlnAsnIleGlnGln--- 1276
 Db 3163 GACATA-----GAGAGGGATATGCTCTGTGAGACCTGCTCATGCCACTGACGACG 3216
 QY 1277 -----SerIleGlnThrGlnAspGlnLeuArgAlaAlaGlnGlnGln 1290
 Db 3217 AACATGCTCAGACAGCCCTCAAAAAAACCTCGGGCTGCTGACGTCTCCAGAGAAAGAA 3276
 QY 1291 LeuArgGlnGlnLysGlnLeuValAspSerPheArgGlnGlnLeuAspLysSerVal 1310
 Db 3277 CTGACCAAGAGAGAGCCCTGATTCAGAACTTCACACAAAGCTA----- 3321
 QY 1311 GlyIleSerSerProAsnHisAspAlaValAlaAsnGlnGlnLysValSerLeuGlnGlu 1330
 Db 3322 -----AACCAAAAGAAA----- 3333
 QY 1331 ValAsnSerLeuGlnSerGlnMetLeuArgGlnArgAspGlnLeuGlnThrSerCys 1350
 Db 1331 ValAsnSerLeuGlnSerGlnMetLeuArgGlnArgAspGlnLeuGlnThrSerCys 1350

Db 3334 -----GAGAGAGTAGAACAGAGAGAGATGAATTAATCAATCAAAATG 3375
 QY 1351 LysAlaLeuValSerGlnLeuGlnLeuLeuArgAlaHisValLysSerValGlnGln 1370
 Db 3376 AGGCAACTA-----GACATGTGATGATCTGCTGCTGAG 3411
 QY 1371 Asn-----LeuGlnIleThrLysLysLeuAsn 1379
 Db 3412 GATCCCCAGAGCTTAAAGACACACCTCACTTCAAAACATTTGGCAAAACCTCGGAA 3471
 QY 1380 GlyLeuGlnLysGlnIleLeuGlnLysSerGlnGlnSerGlnValLeuLysSerMetLeu 1399
 Db 3472 ACACAAAGACAGAGAT-----GAGAGTAGAGAGACCTCTAGACTTCTTG 3519
 QY 1400 GlnAsnLeuLysGlnLysAsnAsnLysLeuLysGlu-----GlnAlaGlnLysTyrSer 1417
 Db 3520 GAACACCTTGAACAAAGCTTAATGAAAGACAGAGAAAGTCAAAATCTGAAATCCCGAGA 3579
 QY 1418 SerLysGlnAsnGlnPheSerLeuGlnGlnValPheSerGlySerGlnLysLeuValAsp 1437
 Db 3580 ATGAGAGAGCAGTTGGCTGAATGAAACCTACGCTGGAAGCTCAGCAGTTAAATAGAG 3639
 QY 1438 GlnIleGlnValLeuLysAlaGlnLeuLysAlaAlaGlnGlnLysValLeuGlnLysAsp 1457
 Db 3640 AAAAAGTGGCTCCTGCAAGGTGCTGATGATTTAAAGACAAAGAGAAACAGTGTAT 3699
 QY 1458 ArgAspTyrPheGlnLeuValGlnThrAlaAsn----- 1468
 Db 3700 CAGAAATCAGCATATATACACGCTGAAGAAATGAACAGAGAAAGTATCAAGAAAGA 3759
 QY 1469 -----ThrAsnLeuValGlnGlnLysLeuGlnThrProLeuGlnAlaAspIleGlu 1485
 Db 3760 CTGCAAAAAGTAATATGTTGAGAAATGCTGAAA-----ATGAAGACAGCTAGAA 3813
 QY 1486 GlnAspSerIleAspArgArgSerGlnGlnMetGlnIleLysValLeuGlnLysLeu 1505
 Db 3814 GAAGTCCAAAGTCCCTTTACAAAGAGATGATGCTTAAGAAATGATGATGATGATGATG 3873
 QY 1506 GlnArgAsnGlnTyrLeuLeuGlnLysArgLeuGlnGlnGlnGlnLysLeuSerAsnLys 1525
 Db 3874 GAACGAAACCAACTTGGAGTCTTAAGCATTCACGAAAGAAAGAAACCACTGATGATCAAG 3933
 QY 1526 LeuGlnIleLeuGlnLysGlnMetGlnThrSerValLeuLysAspLeuGlnGln 1545
 Db 3934 CTGGAAGAAATGATGAAGAAAGAGAGACATCCAGAGATGAGAAATTTAAGAGAG 3993
 QY 1546 LysLeuGlnSerLeuSerGlnAsn-----IleIleLeuLysGlnAsnIleAspThr 1563
 Db 3994 CAGGTGAGTGTCTGCTGAGGAAATGAAAGTTGGTAGTCAACCAAAATTTG----- 4047
 QY 1564 ThrLeuLysHisLysSerAspThrGlnAlaGlnLeuGlnLysThrGlnGlnGlnGln 1583
 Db 4048 -----CATCAAGAGATTCAGTACGTATGGACATCAAGAAAGCAAAAGTCAAGG 4095
 QY 1584 LeuAlaLysAsnLeuAlaIleAlaAlaSerAspAsnGlyProIleThrGlnLysGln 1603
 Db 4096 CTGCTGAGAGACAGAAAGAGTCCGCAAAATGATTTTAAAGAAAGAAAGAAAGA 4155
 QY 1604 ThrSerAla 1606
 Db 4156 AGTGAATCT 4164

RESULT 5

US-09-723-262-1
 : Sequence 1, Application US/09723262
 : Patent No. 6379912
 : GENERAL INFORMATION:
 : APPLICANT: Beraud, Christophe
 : APPLICANT: Sakowicz, Roman
 : TITLE OF INVENTION: No. 6379912el motor proteins and methods for
 : FILE REFERENCE: their use
 : FILE REFERENCE: 1017

```

: CURRENT APPLICATION NUMBER: US/09/723,262
: CURRENT FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: US 09/572,191
: PRIOR FILING DATE: 2000-05-17
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1
: LENGTH: 4757
: TYPE: DNA
: ORGANISM: Human
: US-09-723-262-1

Alignment Scores:
Pred. No.: 1,99e-75 Length: 4757
Score: 1051.00 Matches: 417
Percent Similarity: 42.34% Conservative: 304
Best Local Similarity: 24.49% Mismatches: 548
Query Match: 7.12% Indels: 434
DB: Gaps: 60

US-09-150-867-1 (1-2954) x US-09-723-262-1 (1-4757)

QY 2 SerGluGlyAspAlaValLysValCysValArgValArgProLeuIleGlnArgGlu--- 20
DB 64 AATGAAGGTATGCCATCAAGTTTGTGTGCGAATGCTCTCCCTGCAAGAAAGATCTGGG 123
QY 21 -----GlnGlyAspGlnAlaAsnLeuGlnTrpLysAlaGlyAsnAsnThrIleSerGln 38
DB 124 TCAGCTTATGAGAGAGAG---AACTTATGCTTATGCTGCTGCTCCACAGATCTCCGG 180
QY 39 ValAspLysThr-----LysSerPheAsnPheAsnArgValPheAsnSerHisGlu 55
DB 181 CTGCACTCCAAACCTGAGCCCAAGACCTGACGTTTGATCATGTTGGAGATGGATACC 240
QY 56 SerThrSerGlnIleTyrglnGlnIleAlaValProIleIleArgSerAlaLeuGlnGly 75
DB 241 ACTGAGGAATCTGATTGCGCAACTGTGGCTAAAGACATTTGGAGTCTTGATGAGCGGT 300
QY 76 TyrAsnGlyThrIlePheAlaTyrglyGlnIlePheSerSerGlyLysThrTyrrMetMet 95
DB 301 TATTAATGATGCATCTTGTGCTATGAGACAGACTGCTCAGGAAGACATTTACTATGATG 360
QY 96 Gly-----ThrProAsnSerLeuGlyIleIleProGlnAlaIle 108
DB 361 GGACCATCTGAACTGATTAATTTTCTCATAACTGAGAGAGATTAATCCACGAAGTTT 420
QY 109 GlnGluValPheLysIleIleGlnGluIle-----ProAsnArgGluPhe 123
DB 421 GAATATTGTTTCTTAATTAATGATCGTGAAGAAAGAGCTGAGCTGGAAGAGTTTC 480
QY 124 LeuLeuArgValSerTyrrMetGluIleTyrrAsnGlnIleThrValLysAspLeuGlyAsp 143
DB 481 CTTTGTAAAGTTCCTTATTAATGAATCTACACAGCAGATATGATCTACTG----- 534
QY 144 AspArgThrGlyLysProLeuGlnIleArgGluAspPheAsnArgAsnValTyrrValAla 163
DB 535 GACTCTCATCGGCTGAGCTGACTTAAGGAGCATATCAAGAGAGGAGCTTGTGTGT 594
QY 164 AspLeuThrGluGlnLeuValMetValProGlnHisValIleGlnIleProIleLysGly 183
DB 595 GGTGCGGTGAGCAGAGGTAACCTCAGCTGCTGAAACCTATATCGTGTGTCTGAGAGA 654
QY 184 GlnLysAsnArgHisTyrglyGlnIleThrLysMetAsnAspHisSerSerArgSerHisThr 203
DB 655 TGGAGGAATAGACGTGGGATCAACATCAATGAACAGAGAAATGCTTAGGTCTCAATGCC 714
QY 204 IlePheArgMetIleValGlnSerArgAspArg---AsnAspProThrAsnSerGluAsn 222
DB 715 GCTTTACATTAACAATAGAGTCAATGAGAAAGTAATGAGATTTGCAAT----- 765
QY 223 CysAspGlyAlaValMetValSerHisLeuAsnLeuValAspLeuAlaGlySerGluArg 242
DB 766 -----ATACGACCTCCCTACTCAACCTGCTGATTTAGACAGGATCTGAAGG 813

```

```

QY 243 AlaSerGlnThrGlyAlaGlnGluIleValArgPheLysGlnGlyCysAsnIleAsnArgSer 262
DB 814 CAAAAGATACCATGCCAAGAGGATGATTAAGAGACAGATTAATTAATGATCA 873
QY 263 LeuPheIleGlnGlyGlnValIleLysLysLeuSerAspGlyGlnAlaGly-----Gly 280
DB 874 TTGAGCTGCTGGGCAAGTATTAACAGCACTGTGTGACGTGGGTAATGAAACAGAGA 933
QY 281 PheIleAsnTyrrArgAspSerLysLysLeuThrArgIleLeuGlnAsnSerLeuGlyLys 300
DB 934 CATGTTGCTACAGAGACTCCAACTTAACCTTACTACGCGGATTCCTTGAGGAT 993
QY 301 AlaLysThrValIleIleCysThrIlePheProValSer-----PheAspGlnThrLeu 318
DB 994 GCCAAACAGCCATTAATGCAAAATGTTATCTGATCCAGGTGTTGGGAAACCTTA 1053
QY 319 SerThrLeuGlnPheAlaSerThrAlaLysHisValArgAsnThrProHisValAsnGlu 338
DB 1054 TCAACACTTAACCTTGTCTCAAGAGCCCAAGCTGATTAACAAAGGCAGTAGTAATGA 1113
QY 339 ValLeuAspArgLysAlaLeuLeuLysArgTyrrArgLysGlnIleLeuAspLeuLys 358
DB 1114 -----GACACCCAAAGAAATGTAGCGCTCCAAAGCTGAAGTGAAGAGCTCAAGAA 1167
QY 359 GlnLeuGlnAsnLeuGlnLysSerSerSerGlnThrLysAlaGlnAlaMetAlaLysGlnGlu 378
DB 1168 CACTGGCGGAGCTTGTCTGAGACAGACACCACCAAGAACTTCTCTACAGAGACAA 1227
QY 379 HisThrGlnLeuLeuAlaGlnIleLysGlnLeuHisLysGlnArgLysArgIleTrp 398
DB 1228 AAGAAGACTAATATATGAGTATTTCCAG-----GAAGCAATGTTATTC 1272
QY 399 HisLeuThrAsnIleValAlaIleSerSerGlnGlnLysGlnIleAspGlnArgValLys 418
DB 1273 TTT-----AAGAACTGAAACAGAAAGAG----- 1299
QY 419 ArgLysArgArgValThrTrpAlaProGlyLysIleGlnAsnSerLeuHisAlaSerGly 438
DB 1299 ----- 1299
QY 439 ValSerAspPheAspMetLeuSerArgLeuProGlyAsnPheSerLysAlaLysPhe 458
DB 1300 ---TCTGTGATGAAGAAATTAATCCAAATTAAGAACCTCACCTCAAAAGAAATTT 1356
QY 459 SerAspMetProSerPheProGlnIleAspAspSerValCysThrGlnPheSerAspPhe 478
DB 1356 ----- 1356
QY 479 AspAspAlaLeuSerMetMetAspSerAsnGlyIleAspAlaGlnIleTrpAsnLeuAlaSer 498
DB 1357 -----ATTCATCTTAATTAATGATTAATGATTAATGATTAATGATTAATGATTA 1395
QY 499 LysValThrHisArgGlnLysThrSerLeuHisGlnSerMetIleAspPheGlyGlnIle 518
DB 1396 CAAATTAATACCTTGAAAG-----CTCCAC----- 1422
QY 519 SerAspSerValGlnPheHisAspSerSerLysGlnAsnGlnIleGlnIleProLys 538
DB 1423 -----AAGAAATCCCGGAGAGCTTTCTCCCTGAG 1452
QY 539 AspSerGlyAspMetAlaGlnCysArgLysAlaSerPheGlnLysGlnIleThrSerLeu 558
DB 1453 GAGCAGAGATCTTGTGCTC-----TCAAAATTAAGAAATGAGATTAATGATTAATGAT 1500
QY 559 GlnGlnGlnLeuGlnSerLysGlnGlnGlnLysLysGlnLeuValGlnSerPheGlnLeu 578
DB 1501 CGAAGAACAAATAGAGACACACCCAGAGTTGCAAAAGTATGCTAAGAAATCATTCCTC 1560
QY 579 LysIleAlaGlnLeuGlnGlnGlnLeuSerValLysAlaLysAsnLeuGlnMetValThr 598
DB 1561 AGG-----GAGGAGAAATAGAAAGCTGAGATTA-----TTAAGAGCTGTGAAA 1602

```

```

QY 599 AsnSerArgLuhHisSerIleAsnAlaGluValGlnThrAspValGluLysGluValVal 618
   : : : : :
Db 1603 AGAGCTCAAGAA----- 1614
QY 619 ArgLysGluMetSerValLeuGluLysAspSerGlyTyrAsnAlaSerAsnSerAspLeuGln 638
   : : : : :
Db 1614 ----- 1614
QY 639 AspSerSerValAspLysArgLeuSerSerHisAspLysCysIleGluHisArg 658
   : : : : :
Db 1614 ----- 1614
QY 659 LysMetLeuGluGlnLysIleValAspLeuGluGlnPheIleGluAsnLeuAsnLysLys 678
   : : : : :
Db 1615 ---ATGATGCCCAAGACATTCGAAACAGAAAGCTTCTGTAATTAAGTGCATG 1671
QY 679 SerGluAsnAspLysGlnLysSerSerGluGlnAspPheMetGluSerIleGluLeuCys 698
   : : : : :
Db 1672 GAGAAAGTGACAAA-----ATCACAGAGATTTTCACCTTAAGCTCAGAAAGAG 1722
QY 699 GluAlaIleMetAlaGluLysAlaAsnAlaLeuGluGluLeuAlaLeuMetArgAspAsn 718
   : : : : :
Db 1723 CCATGTTGTTTGA----- 1737
QY 719 PheAspAsnIleIleLeuGluAsnGluThrLeuLysArgGluIleAlaAspLeuGluArg 738
   : : : : :
Db 1738 ----- 1779
QY 739 SerLeuLysGluAsn---GlnGluThrAsnGluPheGluIleLeuGluLysGluThrGln 757
   : : : : :
Db 1780 GAGCTGAATTAATTCAAAGCAAGATATGAGAAATTC-----AAGAACTTACT 1827
QY 758 LysGluHisGluAlaGluLeuIleHisGluIleGlySerLeuLysLysLeuValGluAsn 777
   : : : : :
Db 1828 AGGAAAGGACAGTAGAATTGGAATCAGACCTTCAGTCTTGCAAAAGCG----- 1878
QY 778 AlaGluMetTyrAsnGlnAsnLeuGluGluAspLeuGluThrLysThrLysLeuLeuLys 797
   : : : : :
Db 1879 -----AACCTTAATCTTGAAACCTTTTGAGACA---ACAAAAGCCTGAG 1923
QY 798 GluGlnGluIleGlnLeuAlaGluLeuArgLys---ArgAlaAspAsnLeuGln----- 814
   : : : : :
Db 1924 CGGCAAGAAAGTT-----TTCAGCTGAATTAATTCATCTGAACACTTAAGATTATA 1977
QY 815 -----LysLysValArgAsnPhe-----AspLeuSerValSer 825
   : : : : :
Db 1978 ACTACACCAACCAAGGCTACCAACTCATTCGCCGACCAAGTACCAAAATTAAGCCCGAA 2037
QY 826 MetGluAspSerGluLysLeuCysGluGluIlePheGlnLeuLysGlnSerLeuSerAsp 845
   : : : : :
Db 2038 ATGGGAAGCTTTGGCTCTCTA-----TACACTCAGAAATTCAGCATATTTAGAT 2085
QY 846 AlaGluAlaValThrArgAspAlaGlnLysGluCysSerPheLeuArgSerGluAsnLeu 865
   : : : : :
Db 2086 AATATATTAATTAATGAGCCAGTTCTCTGAG----- 2118
QY 866 GluLeuLysGluLysMetGluAspThrSerAsnTyrAsnGlnLysGluLysAlaAla 885
   : : : : :
Db 2119 -----ATGAATGAACAAGCTTTT 2136
QY 886 SerLeuPheGluLysGlnLeuGluThrGluLysSerAsnTyrLysLysMetGluAlaAsp 905
   : : : : :
Db 2137 GAGGCCATTTCTGAGAGCTTGAACAGTGACAGACAAATGACTGCTCTTCAGAAACAA 2196
QY 906 LeuGlnLysGlu-----LeuGlnSerAlaPheAsnGluIleAsnTyr 919
   : : : : :
Db 2197 CTGATGATGAAGAGACATAAAACCTTAAGCTTCAGCAGCATGTGTGACAAACCTGGAACAT 2256
QY 920 LeuAsnGlyLeuLeuAlaGlyLysValProArgAspLeuLeuSerArgValGluLeuGlu 939
   : : : : :
Db 2257 CATCTCAACCAATG-----CAGAGAGCTTTCTCATCAGAAAGAAAT--- 2298
QY 940 LysLysValSerGluPheSerLysGlnLeuGluLysAlaLeuGluLysAsnAlaLeu 959

```

```

Db 2299 -----GATTGACCAACAGCAGAGAGAGCTTCTCTCAGACTTAAGTGCCTT 2346
QY 960 GluAsnGluVal---ThrCysLeuSerGluTyrLysPheLeuProAsnGluValGluCys 978
   : : : : :
Db 2347 GAAAGCAGAGCTTCAGAGACATCAAAATGAAATGCTTTTGAAGAGAGCTACATGAC 2406
QY 979 LeuLysAsnGlnIleSerLysAlaSerGluGluIleMetLeuLeuLysGluGlu----- 996
   : : : : :
Db 2407 CTGCGAGTAGTCCCTTCATCTGCTGACAGAGAGCTTCTTCAGTGAATTTGGAATATAGT 2466
QY 997 -----GlyGluHisSerAlaSerIleIleSerLysGlnGluIleMet 1011
   : : : : :
Db 2467 TCATTCAAAACGAATCAGAGAGAAAGAAATTCAAACAACTTCGAAAGACATGATGTA 2526
QY 1012 GlnGluGlnSerGluGlnIleLeuGlnLeuPheThrAspGluValThrHisThrGlnSerLys 1031
   : : : : :
Db 2527 CAGCTTCATTTAGATTAATCTCAGGTTAGAAAAGAAAAGCTGCTGAGACAAAGCCTGC 2586
QY 1032 ValGlnGlnThrGluGluGlnTyrLeuGluMetLysLysMetHisAspAspLeuPheGlu 1051
   : : : : :
Db 2587 CTACAGATTCCTATGACACACTTACAGAAATTAATGAATTTGAGATTGACCAACTTCA 2646
QY 1052 LysTyrIleArgAsn---LysSerGluAlaGluAspLeuLeuArgLysGluMetGluAsnLeu 1070
   : : : : :
Db 2647 AGAAACCTCCAAACCTTCAAAAGAAATGAAACTGTGAAATCTGATCTGAATTAATTTG 2706
QY 1071 LysGlyThrMetGluSerValGluValLys-----IleAla 1082
   : : : : :
Db 2707 ATGGACCTTCCTGAGCAGAAAAGAACCCATTAACAAATTTATCATTCAGTTGAGAGA 2766
QY 1083 AspThrLysHisGluLeuGluGluThrIleArgAspLysGluGlnLeuLeuHisGluLys 1102
   : : : : :
Db 2767 GATTAAGAAACAGCTTCTTAAGAAATCTTAAGAAATCTTGAGCGCTGCTGCTGAGAGAAA 2826
QY 1103 ---LysTyrPhePheGlnIleMetGlnThrIlePheProIleThrProLeuSerAspSer 1121
   : : : : :
Db 2827 CAGAAAGACAGCGCCAAAGTGAGAGCAGCAGATGGCAAAAGTACAGAAACTTAAGAGAGC 2886
QY 1122 LeuProProSerLysLeuValGluLysAsnSerGlnAspProIleGluIleAsnAspTyr 1141
   : : : : :
Db 2887 TTG----- 2889
QY 1142 HisAsnLeuIleAlaLeuAlaThrGluArgAsnAsnIleMetValCysLeuGluThrGlu 1161
   : : : : :
Db 2890 -----CTTCTACTGAAAAA---GTGATGAGTTCCTCGAAAAAGTCT 2928
QY 1162 ArgAsnSerLeuLysGluGlnValIleAspLeuAsnThrGlnLeuGlnSerLeuGlnAla 1181
   : : : : :
Db 2929 AGAGATTCTGATGAAGAAAGTGTAGCTGACCTCATGACAGATCCAGAGCTAAGATCA 2988
QY 1182 GlnSerIleGluLysSerAsp---LeuGlnLysProLysGlnAspLeuGluGluGlu 1200
   : : : : :
Db 2989 TCGGTCTGTGAGAAACAGAAACTATAGACACCCTGAAACAGAACTGAAAGACATTAAT 3048
QY 1201 ValLys-----LeuLeuGluGluMetGluLeuLeuLysGlyIleThrAspSer 1217
   : : : : :
Db 3049 TGCAAATACACTCTGCTTGTGTGACAGAGAA-----GAGAGC 3087
QY 1218 GlnLeuSerIleGluLysLeuGlnLeuGluAsnLeuGluValThrGluLysLeuGlnThr 1237
   : : : : :
Db 3088 AGAGTGTGTATCAAGAAAGAGATGATATCTGATGTAAGAAACCCCTTAGG--- 3144
QY 1238 LeuGlnGluGluMetLysAsnIleThrIleGluArgAsnGluLeuGlnThrAsnPheGlu 1257
   : : : : :
Db 3145 -----CTGAGATTAATCTTTTGTGAG 3162
QY 1258 AspLeuLysAlaGluHisAspSerLeuLysGlnAspLeuSerGluAsnIleGluGln--- 1276
   : : : : :
Db 3163 GACATA-----GAGAGGATATGCTCTGTGAGAGACTGCTCATGCCACTGAGCAGCTG 3216
QY 1277 -----SerIleGluThrGlnAspGluLeuArgAlaAlaGlnGluGlu 1290

```

Db 3217 AACATGCTCACAGAGGCGCTCAAAAAACATCGGGGCGTGCAGTCTGCCAGAGAGAA 3276
 QY 1291 LeuArgGluGlnGlnLeuValAspSerPheArgGlnGlnLeuLeuAspCysSerVal 1310
 Db 3277 CTGACCAAGAAAGAACCCCTATTCAGAACTTCAGCAACGTA----- 3321
 QY 1311 GlyIleSerSerProAsnHisAspAlaValAlaAsnGlnGlnValSerLeuGlyGlu 1330
 Db 3322 -----AACCAAAAGAAA----- 3333
 QY 1331 ValAsnSerLeuGlnSerGluMetLeuArgGlyGluArgAspGluLeuGlnThrSerCys 1350
 Db 3334 -----GAGAGATGACACAGAGAAAGAAATGATATTAACCTCAAAATG 3375
 QY 1351 LysAlaLeuValSerGluLeuGlnLeuValArgAlaHisValysSerValGlyGly 1370
 Db 3376 AGGCAACTA-----GAACTGTGATGGATTCGCTGCTGAG 3411
 QY 1371 Asn-----LeuGluIleThrLysLysLeuAsn 1379
 Db 3412 GATCCCGAGATCCTTAAGACACACACTCTTCAACACATTGGCAAAACTCCTGGAA 3471
 QY 1380 GlyLeuGlnGlyGlyIleLeuGlyLysSerGluLeuSerGluValLeuLysSerMetLeu 1399
 Db 3472 ACACAAAGAACAGAGATA-----GAGATGGAGAGAGCCCTTAAGACTTCTTG 3519
 QY 1400 GluAsnLeuGlyGluAspAsnLysLeuGlyGlu-----GlnAlaGluGlyTyrSer 1417
 Db 3520 GAACACCTTGTAAACAAAGCTTAATGAAGACAGACAGAAAGCAAAATGCTGAAATCCTCGA 3579
 QY 1418 SerLysGluAsnGlnPheSerLeuGluValPheSerGlySerGlnLysLeuValAsp 1437
 Db 3580 ATGAAGAGACAGTGTGCTGAATAGAAACCTACGCTGAAAGCTGAAAGCTGACATTAATGAG 3639
 QY 1438 GluIleGluValLeuLysAlaGlnLeuLysAlaAlaGluGluValGlyLeuGluIleLysAsp 1457
 Db 3640 AAAAATGCGCTCGTCAAGCTCAGCTGATGATTAATAAGACAAAGAAAGAAACAGTGTAT 3699
 QY 1458 ArgAspTyrPheGluLeuValGlnThrAlaAsn----- 1468
 Db 3700 CAGAAATATCCAGATATCAACACAGCTGAGATGACAAAGAAAGATCAAAAGAAAGA 3739
 QY 1469 -----ThrAsnLeuValGluGlyLysLeuGlnThrProLeuGlnAlaAspHisGlu 1485
 Db 3760 CTTCGCAAAAGATAAATAGTTGAAGAAATGCTGAAA-----ATGAAGACAGACTAGAA 3813
 QY 1486 GluAspSerIleAspArgAspSerGluGluMetGluIleLysValLeuGlyGluLysLeu 1505
 Db 3814 GAACTCCAAAGTCCCTTTACACAAAGAGATGGAATGCTTAGAATGACTGATGAAGTC 3873
 QY 1506 GluArgAsnGlnTyrLeuLeuGlnArgLeuGlnGlnGlnLysLeuGlnLeuSerAsnLys 1525
 Db 3874 GAACGAACCCAAACTTGGAGCTCAAAAGCATTCACAGAAAAAGAAAGCACTGATCAAG 3933
 QY 1526 LeuGluIleLeuGlnLysGluMetGluThrSerValLeuLysAspAspLeuGlnGln 1545
 Db 3934 CTGGAAGAAATGTATGAAGAAAGAGAGAACATCCAGAGATGGAATGTAAAGAGAG 3993
 QY 1546 LysLeuGlnSerLeuLeuSerGluAsn-----IleIleLeuLysGluAsnIleAspThr 1563
 Db 3994 CAGGTGAGTGTCTTGTGAGAAAGAAATGCTGATCCCAAAATTTG----- 4047
 QY 1564 ThrLeuLysHisHisSerAspThrGlnAlaGlnLeuGlnLysThrGlnGlnLeuGln 1583
 Db 4048 -----CATCGAAGATTCAGTACGTAGTGCATTAAGAAAGAAAGAAAGTACAGG 4095
 QY 1584 LeuAlaLysAsnLeuAlaIleAlaAlaSerAspAsnCysProIleThrGlnGlnLysGlu 1603
 Db 4096 CTTCGCTAGAGAGACAGAAAGTTGCGTCCGCAAAATGTATTTTAAAGAAAGAAAGAAAGA 4155
 QY 1604 ThrSerIle 1606
 Db 4156 AGTGAATCT 4164

RESULT 6
 US-09-723-219-1
 : Sequence 1, Application US/09723219
 : Patent No. 6391613
 : GENERAL INFORMATION:
 : APPLICANT: Beraud, Christophe
 : APPLICANT: Sakowicz, Roman
 : APPLICANT: Mood, Kenneth
 : TITLE OF INVENTION: No. 6391613el "motor proteins and methods for
 : their use
 : FILE REFERENCE: 1017
 : CURRENT APPLICATION NUMBER: US/09/723, 219
 : PRIOR FILING DATE: 2000-11-27
 : PRIOR FILING DATE: US 09/572,191
 : NUMBER OF SEQ ID NOS: 6
 : SOFTWARE: FASTSEQ for Windows Version 4.0
 : SEQ ID NO 1
 : LENGTH: 4757
 : TYPE: DNA
 : ORGANISM: Human
 US-09-723-219-1
 Alignment Scores:
 Pred. No.: 1,99e-75 Length: 4757
 Score: 1051.00 Matches: 417
 Percent Similarity: 42.34% Conservative: 304
 Best Local Similarity: 24.49% Mismatches: 548
 Query Match: 7.12% Indels: 434
 Gaps: 60
 US-09-150-867-1 (1-2954) x US-09-723-219-1 (1-4757)
 QY 2 SerGluGlyAspAlaValLysValCysValArgValArgProLeuIleGlnArgGlu--- 20
 Db 64 AATGAAGTGAATGGCAATCAAAAGTTTGTGCGAAATTCCTCTCGCAAAAGATCTGGG 123
 QY 21 -----GlnGlyAspGlnAlaAsnLeuGlnIleThrLysAlaGluAsnAsnThrIleSerGln 38
 Db 124 TCAGCTGATGAGAGACAG---AACTTATGCTTATCTGTGCTGCTCCACAGATCTGGG 180
 QY 39 ValAspGlyThr-----LysSerPheAsnPheAspArgValPheAsnSerHisGlu 55
 Db 181 CTGCACCTCCAAACCTGAGCCCAAGACCTTCACGTTGATGATGTGAGATGGATGATACC 240
 QY 56 SerThrSerGlnIleTyrGlnGlnIleAlaValProIleIleArgSerAlaLeuGlnGly 75
 Db 241 ACTCAGGAATCTGTATTCGCAACTGTGCTAAAGCAATTTGTGAGATCTTCGATGACGCGT 300
 QY 76 TyrAsnGlyThrIlePheAlaTyrGlnThrSerSerGlyLysThrTyrThrMetMet 95
 Db 301 TATATGTACCATTTTTCATATGACAGACTGCTCAGGAAAGACATTTACTATATAGT 360
 QY 96 Gly-----ThrProAsnSerLeuGlnIleIleProGlnAlaIle 108
 Db 361 GGACCACTGAACTGATTAATTTTTCATTAACCTGAGAGAGATTAATCCCAAGATTTT 420
 QY 109 GlnGluValPheLysIleIleGlnGluIle-----ProAsnArgGluPhe 123
 Db 421 GAATATTTGTTTCTTAATGATCGTGAAGAAAGAGCTGAGCTGGAAGAGTTTC 480
 QY 124 LeuLeuArgValSerTyrMetGluIleTyrAsnGlnThrValLysAspLeuLeuCysAsp 143
 Db 481 CTTTGTAAGTCTCTTATATGAATCTACACAGACAGATATGTGATCTGCTG----- 534
 QY 144 AspArgArgLysLysProLeuGluIleArgGluAspPheAsnArgAsnValTyrValAla 163
 Db 535 GACTCTGCATCGCGCTGAGCTGACTTACTTAAGAGACATTAAGAAGAGAGCTTGTGTT 594
 QY 164 AspLeuThrGluGluLeuValMetValProGlnHisValIleGlnIleThrLysLysGly 183
 Db 595 GGTGCGGTGAGCAGGTGTGAACCTCAGCTGCAACCTATACAGTGTGTCTGAGAGA 654

Db 1279 -----CAG 1281
 QY 511 Sermetileasppheglycylleiseraspservalglnphehisaspserserlyscu 530
 Db 1282 TCCCTGGTAGAG-----GAGAAATGAA 1302
 QY 531 Asnleuglnlyrleuprollysaspsersglyasmetaglucysarglyslaser 550
 Db 1303 AAATTAACTGTGTCTGAGCAAGCGAGCTGTGACAGACCCAGATGTTGGAGGAGATC 1362
 QY 551 Phegllysgluilethrserleuglnleuglnserlysgluuglnlysls 570
 Db 1363 ATTTGACAGAG-----CAAGTGAAATGAAACGCAACGCCAAGCTGAGAGCTCAGG 1416
 QY 571 Glu-----leuvalglnserphegluleulysile 580
 Db 1417 CAGATGGCGGCTGCAMCTGATCTTCAAAAGCTAGTGAGACTTTGGAGACCCAG--- 1473
 QY 581 Alaagluleuglnleuglnleuservalylsalaalsnleuglnmetvalthranser 600
 Db 1474 ---GATTGAAAGAAATGTAGATATTGTAACTGCGACCACTGATTAACCACTTA 1530
 QY 601 Arggluhisserie-----Asnaglulvalglnthraspvalglulysglulval 617
 Db 1531 TCAGATGAACCTGTGCTTGCACGCGTGCACCATTTGACTGCGGTGAGAAAGAGAA--- 1587
 QY 618 Valarglysglumetservalleuglnlyaspsersglytyrasnlaaseraspsersleu 637
 Db 1588 -----GCTCAAGTGGAAACCACT 1605
 QY 638 Gluaspserservalaspglyllyargleuseraserhisaspglucylleuglnhis 657
 Db 1606 CCAAGACACACAGCTCTTCTGACGCTTTACCACTCAGCAT-----GCTCTCAT 1656
 QY 658 Arglysmetleuglnlyslsilevalaspleuglnlupheile-----glu 673
 Db 1657 CAACCTGAGATGCTAAGAGAGGTGTTGATGAATAAGCCCTTGCACTGAAAGAGGCC 1716
 QY 674 Asnleuasnlylsersercluasnasplysglnlysserserlysglnaspheketlu 693
 Db 1717 CTACTTGAAGAGATGACTCAACAGCACCAACCA-----CTACAG 1755
 QY 694 Serlleuglnleucysglualallemetaglulysalaasnalauglnleuula 713
 Db 1756 CCCATTCAGTT----- 1767
 QY 714 Leumetaraspasnpheaspasnilleleuglnusnglnthrleulysarglulle 733
 Db 1768 CAATACACAGATTAATAAATACTAGAAATTAGCAATCAATCTGCAAAAGAGAAAG 1827
 QY 734 Alaaspleuglnlyargsetleu-----lysgluasnnglnluthrasnglnlupheilu 750
 Db 1828 GAAATATTTGGTCTGAACTTCAACAGCAAGCAAGAAAGAAATGCCAACCAAGCTGAGT 1887
 QY 751 lleuglnlysglnthrnglnlysluhisglualaglnleu----- 764
 Db 1888 GAGCAGCGGTCCAAACTTCTCAGAGAGGTGAGAGGTCAAAATAGCTGATCAAGAGAA 1947
 QY 765 llehisglulleglyserleuylsleuvalgluasnalaglumettyrasnglnasn 784
 Db 1948 CTGAATAGCAGCTCAAACTTCAAACTAAGCAATCCACAGAGCGTACTGTCCAAA 2007
 QY 785 leuglnleuglnaspleuglnluthrlystrhrleuileuylsnglnleuglnleula 804
 Db 2008 CTGACCAAGGAGATA-----CGGATGATGAAACCAAGCGGGTACAGTTA--- 2052
 QY 805 Gluubuarlyysargalaspasnleuglnlyslsvalargasnpheaspleuserval 824
 Db 2053 ---ATGGGTCAAATGAAAGAGATGCTGAG---AAGTTTACACAGTGS----- 2094
 QY 825 Sermetcllyaspsersglyllyllyleuglnlullepheglnleuylsnglnserleuser 844

Db 2095 -----AACGAGAAAGAGACAAAGAAATATACAGTTAAAGAA----- 2133
 QY 845 Aspalaglualavalthrargaspalaglnlysglucyserspheleuargsetluasn 864
 Db 2134 -----CGAGACCGTAAAGAGCAAA-----TATACGTG 2160
 QY 865 leuglnleuylsglulysmetglulaspthrserantptyrasnnglnleuglnleula 884
 Db 2161 CTGAATCTTGAAGAAACTCTCCAGAAACATCAATGCTGCTCAGACGTAACGAGAGNG 2220
 QY 885 Alaserleupheglulysgluleuglnluthrnglnlyssersasnlyrlyslasmetglula 904
 Db 2221 GCAGCAGCTGCCAACACAGCGTCTCAAGGATGCTCCAGAAACACGGGAGGTT---GCA 2277
 QY 905 Aspleuglnlysglnleuglnserlaphasnglnlleasnlyrleuasnnglnleu--- 923
 Db 2278 GATTAAGCGGAAAGAGACACACAGC-----CGTGAAATGGA 2313
 QY 924 ---leuallaglylyvalproargaspheuserasrargvalgluleuglnlyslsval 942
 Db 2314 GGCACCTGCAGCTCGAGTAGAGAAATGCTTGAATTAATTAAGCTTATGCTGACT 2373
 QY 943 Serglupheserlysglnleuglnlyslsalauglnlulysasnalaleuglnlunlu 962
 Db 2374 GAGGAAGCCAAACGCCATCTGAATGACCTCTGAAAGACAGAAAGATCTGCTCAGGAT 2433
 QY 963 Valthrlycysleuserglutyrlyrlyspheleuproasnnglnlucylleulysngln 982
 Db 2434 GTGTTCAACTTAAGAA----- 2451
 QY 983 lleserlysalaserglulullemetleuylsnglnlucyluhiseralaser 1002
 Db 2452 -----AAAAAGAAATCGGAGAAATCCACCTCT 2481
 QY 1003 llelleserlysglnlulleleuclnglnlunserglulnleleuglnleuthr 1022
 Db 2482 AAATCCCGAAGTAGATCTCCCTTCTGAGGTCATGTCAGTATTTGGACAGAA 2541
 QY 1023 Aspglulvalthrhisclnserlyslsvalnglnluthrnglnlucylnttyrleuglnmet 1042
 Db 2542 GATTGTATTACA-----AAACAGATTGAACCTTAGACCTGAATGGAATGCACTC 2589
 QY 1043 lyslysmethls---Aspaspleupheglululyslyrlyleargaspnlaserclulaglu 1061
 Db 2590 AGGAGCTGCAGATGCTGACCTACACAGAAAGCTGCGATGCGAAGATGAAATGAG 2649
 QY 1062 Aspleuuararglumetcluasnleuylselythrmetgluservalglulvalysile 1081
 Db 2650 CG-BCAAAACAATGCTGGAGAAATATTCACACCATTCGGAAGCC-----AAGTGT 2699
 QY 1082 Alaaspthrlyshisgluleuglnluthrleargaspnyglulnleuulhaglul 1101
 Db 2700 GGC---CTGAATATTTGATTGAGAGAGCTGCTCTCTCAAA-----ATACAT--- 2744
 QY 1102 lyslystyrpheheglinalametcnlthrlepheproilthrproleuseraspsers 1121
 Db 2745 -----GTCACCAAACTTGAAGAAACGCG 2765
 QY 1122 leuproposerlyslsleuvalglulysnsergluaspproilleuglnleasnsplyr 1141
 Db 2766 CTGACACAGACAGCGCAGCTGTGCTGACATGCAAGAGATGTA---TTTGAGAACAA 2822
 QY 1142 Hisasnleuilealaleuathrcglurgrasnnglnleuclnleuclnleuglnluthr 1161
 Db 2823 AATCTTTTCTGAGATGACACAGATTTACAGCTGAGCTGCTGCTGACATGAGCAACAG 2882
 QY 1162 Argasnserleuylsnglnlvalilleaspleuasnthrleuglnleuglnleucln--- 1180
 Db 2883 CAC-----CAAGAGAAAGGTGTATACCTTGACACAGCTGACAGAAAGCAAAATG 2933
 QY 1181 ---Alaglnserlleuglnlyser----- 1187
 Db 2934 GCAGAGAGCAGTTAGAGAAATCAGCCAGTGAAGAAAGCAACAGATGCTGAGCACACTG 2993

QY 1188 -----AspleuGlInuysProLySGInAspleuGlInuGlInuValLys 1202
 Db 2994 CAGGTGACAGATGAGAACTTGAGAAAGATGCGAGAGTGTGGAGAAATACAG---CAG 3050
 QY 1203 LeuLeuLeuGlInuMetGlInuLeuLysGlyHisLeuThraAspSerGlnLeuSerIleGlu 1222
 Db 3051 CTTCTCCAGAGATGCAATCATCAAGCAAGAACTGATCTCTCCAGGATGACAGCAGCA 3110
 QY 1223 LysLeuGlInuLysn-----LeuGlInuValThrcGluLysLeuInThrcLeuGln 1239
 Db 3111 CAGAAACATCTTCTTATATATACCTTCTTATCTCCAGACTCTTCTTTAAATATATCCA 3170
 QY 1240 GluInuMetLysAsnIleThrIleGluArgAsn-----GluLeuGlnThraAspGlu 1257
 Db 3171 CCAAGACCAAACTCTGCTGTAAAGAAAGTTCTGAGCAAGATGACATGACAG 3230
 QY 1258 AspleuLys-----AlaGlnHisAspSerLeuLysGlnAspleuSerGlnIleGlu 1275
 Db 3231 GATCTAAATATATGTTTCAAGATCTGTGTGAATGACATGAAAGATGTGATGATGCG 3290
 QY 1276 GlnSerIleGluThrcGlnAspGlu-----LeuArgAlaIleGln 1288
 Db 3291 GACATGATGAGGGGATGATGAGATGAGATGAGACCAACAAATTTACTGATGATGCG 3350
 QY 1289 GluInuLysn-----GluGlnLysGlnLeuValAspSer 1300
 Db 3351 AGAACATCCAAAGGATGCTCTGCAAGGGGTGTGGAGCAAGACAGTGT-----GGG 3404
 QY 1301 PheArgGlnLeuLeuAspCysSerValGlyIleSer-----SerProAsnHisAspAla 1319
 Db 3405 TCGAGAACCAAAAGTCAACACTGTGTGTGACTGTGTGACCCACCAAAAGTGTGCG 3464
 QY 1320 ValAlaAsnGlnLysValSerLeuGlyGluValAsnSerLeuGln 1335
 Db 3465 AACCCGACCAAGGCAAGATGACTGTGGCAGCTGTGAACAGCCAG 3512
 RESULT 8
 US-09-722-139-1
 ; Sequence 1, Application US/09722139
 ; Patent No. 6355471
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; APPLICANT: Freeman, Richard
 ; TITLE OF INVENTION: NO. 6355471el motor proteins and methods for
 ; FILE REFERENCE: 1055
 ; CURRENT APPLICATION NUMBER: US/09/722,139
 ; CURRENT FILING DATE: 2000-11-24
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 4176
 ; TYPE: DNA
 ; ORGANISM: Human
 US-09-722-139-1
 Alignment Scores:
 Pred. No.: 1,65e-64 Length: 4176
 Score: 915.50 Matches: 416
 Percent Similarity: 40.59% Conservative: 266
 Best Local Similarity: 25.51% Mismatches: 536
 Query Match: 6.20% Indels: 433
 Gaps: 71
 US-09-150-867-1 (1-2954) x US-09-722-139-1 (1-4176)
 QY 6 AlaValLysValCysValArgValArgProLeuIleGlnArgGluGlnLysGlnAla 25
 Db 7 TCGGTGAGGTGGCGGTGAGGGTCCGGCCCATGATCGAGGAAAGCATCTTGAGAGCC 66
 QY 26 Asn-----LeuGlnTrpLysAlaGlyAsnAspThrIleSerGlnVal----- 39

Db 67 AAGTCATTTATTCAGATGAGAAAGCAAAACGACATCATCAAACTTAAGATACCAGA 126
 QY 40 -----AspGlyThrLysSerPheAsnPheAspArgValPhe 51
 Db 127 GGAGGCACTGGGAGACTCAGAGAGACGACGACCAAGACCTTACCTATGACTTTCTTT 186
 QY 52 AsnSerHisGluSerThrSer-----GlnIleArgGlnIle 64
 Db 187 TATTCGTGATACAAAAGCCAGATTAACGTTTACAGAAATGGTTTCAAAACCTTC 246
 QY 65 AlaValProIleIleArgSerAlaLeuGlnGlyTrpAsnGlyThrIlePheAlaTrpGly 84
 Db 247 GGCACAGATGTGTGAGATGCTGATTTGAGGTATTAAGCTGTGTCTTCCATATGG 306
 QY 85 GlnTrpSerSerGlyLysThrTrpThrMetGlyThrProAsnSerLeuGlyIle 104
 Db 307 CAACATGATCTGAAAGATCATCACTATGATGAGAAATTCGAGATTCGCTTAATA 366
 QY 105 ProGlnAlaIleGlnGluValPheLysIleIleGlnGluIle-----ProAsnArg 121
 Db 367 CCTCGATCTGGAAGGACTCTTCAGTCGATTAATGAAACCAACAGATGGATAGACT 426
 QY 122 GluPheLeuLeuArgValSerTrpMetGluIleTrpAsnGluThrValLysAspleuLeu 141
 Db 427 TCTTTTCCAGCTGAACTGACCTACTTAAATTTATACGAACGCTGTGAGATCTACT 486
 QY 142 CysAspAspArgArgLysLysPro-----LeuGluIleArgGlnAspPheAsn 157
 Db 487 -----CGGCGGAAGTCAATCAAACTTCAATTGAGAGCTCGGTGACATCCAAA 537
 QY 158 ArgAsnValTrpValAlaAspLeuThrGluLeuValMetValProGlnHisValIle 177
 Db 538 GAAGGCCCTTATGATGAGATGATTCCAAAATTTAGTACAGAAATTAAGTACAGAA 597
 QY 178 GlnTrpIleLysLysGlyLysAsnArgHisLysTrpGlyLysMetAsnAspHis 197
 Db 598 GAACCTATGATGCGGAGATATCAACCGACACCGACGACGATGGATGAAACGACGTC 657
 QY 198 SerSerArgSerHisThrIlePhe-----ArgMetIleValGluSerArgAspArgAsn 215
 Db 658 ACTAGACGCTCATAGCCCTTCCTCACCATCAAGTTCACCTACGCTAATTTACTGTGA 717
 QY 216 AspProThrAsnSerGlnLysnAspGlyAlaValMetValSerHisLeuAsnLeuVal 235
 Db 718 ARGCCA-----TGTGAA-----ACGTCATGATGATTCACCTTGGTT 753
 QY 236 AspleuAlaGlySerGluArgAlaSerGlnThrGlyAlaGluGlyValArgLeuLysGlu 255
 Db 754 GATCTGCGCGGAAGTGAAGGTGAGATGCGACGAGCCAGCGGGGTGAGCTAAAGGAA 813
 QY 256 GlyCysAsnIleAsnArgSerLeuPheIleLeuGlyGlnValIleLysLysLeuSerAsp 275
 Db 814 GGGGGAATATTTAAACAAGTCCCTTGACTGTGGGGAAGCTCATTTCTCCCTTACGTAT 873
 QY 276 GlyGlnAlaGlyGly-----PheIleAsnTrpArg 285
 Db 874 TTATCTCAGAGATGCTCAAAATCTTGCAGAAAGACCAAGTTCCTGCTTCACAG 933
 QY 286 AspSerLysLeuThrArgIleLeuGlnAsnSerLeuGlyLysAlaLysThrValIle 305
 Db 934 GATCTGCTTGACTGTGTTTAAAGATAGCTTGAGGAAACCTTAACCTATACAG 993
 QY 306 IleCysThrIleThrPro-----ValSerPheAspGluThrLeuSerThrLeuGlnPhe 323
 Db 994 ATTGCCACCATTTACCTCTGCTGATCATTTATGAGAAACCTTAAGTACTTCTCCAT 1053
 QY 324 AlaSerThrAlaLysHisValArgSerThrProHisValAsnGluValLeuAspAspGlu 343
 Db 1054 GCAAATGAGCCCAAAACATCATCAACCAAGCTTACCATTAATGAGATGCCAAGCTCAAA 1113
 QY 344 AlaLeuLeuLysArgTrpTrpGlyGluIleLeuAspLeuLys----- 357
 Db 1114 ----CTTATCCGTGAGCTGAGCTGAATATGCAAGCTGAAACAGCGTGTGCTCAAGG 1170

QY 358 LysGlnLeuGluAsnLeuGlnSerSerGlnThrLysAlaGlnAlaMetAlaLysGlu 377
 DB 1171 AATAGATGCTCCCTCTTACACTCCCA-----GCTTAAGTATGAG 1215
 QY 378 GluHisThrGlnLeu-----LeuAlaGluIleLysGlnLeuHisLysGluArgGlnAspArg 396
 DB 1216 GAAAAACTTCGACAGATGAAGCAAGATTCAAGATTGACCAAGAAATGACAAATAG 1275
 QY 397 IleThrHisLeuThrAsnIleValValAlaSerSerGlnGlnSerGlnGlnAspGlnArg 416
 DB 1276 ---TGGAATGAAACC-----CAAAAT 1293
 QY 417 ValLysArgLysArgArgValThrTrpAlaProGlyLysIleGlnAsnSerLeuHisAla 436
 DB 1294 ATTTGAAAGACAAACTCTA-----GCCCTGACGAAA 1326
 QY 437 SerGlyValSerAspMetLeuSerArgLeuProGlyAsnPheSerLysAla 456
 DB 1327 GAAGGATTTGA-----GTTCTT 1344
 QY 457 LysPheSerAspMetProSerPheProGlyIleAspAspSerValCysThr----- 473
 DB 1345 TTGGATTCTGAACGCTCATTTTGATTGATTCATGATGACCTTTTGAGACTGGAATC 1404
 QY 474 -----GluPheSerAspPheAspAlaLeuSerMet 484
 DB 1405 ATCTATATCATTTAAAGGAAGCTCAGACATACCTGTGTAGACGATGCTCCACGGAG 1464
 QY 485 MetAsp-----SerAsnGlyIleAspAlaGlu-----TrpAsnLeu 496
 DB 1465 CAGATATTTGTTCTTCTTACCTGCTGACTGTGAGAGTACATTCATCTTGAATAATATC 1524
 QY 497 AlaSerLysValThrHisArgGlyLysThrSerLeuHisGlnSerMetIleAspPheGly 516
 DB 1525 GGGGGGACAGTACT---CTGATACCCCTGAGTGGTCCACAGTCTGTGCAATGGTGT 1581
 QY 517 GlnIleSerAspSerValGlnPheHisAspSer-----LysGlnAsn 531
 DB 1582 CAGATCTGTGAGGCGCACACATCTAAATCAAGTCTGTGATTCCTTGGGAGAACCAAT 1641
 QY 532 GlnLeuGlnTyr---LeuProLysAspSerGlyAspMetAlaGlySarGlyS----- 548
 DB 1642 ATGTTCGCTTTAACCTTCAACCAAGGAAGCGCCAAAGCTCAGGAGAGAAAGATGGC 1701
 QY 549 -----AlaSerPheGlnLysGluIleThrSerLeuGlnGlnGlnSerLysGlu 566
 DB 1702 CTTCGTCTCTCTTACGCTTGTCCATGACCGACTC-----TCGAAGTCC 1746
 QY 567 GluGlnLysLysGlnLeuValGln-----SerPheGluLeuLysIle 580
 DB 1747 CGTGAGAACTGTCTGACATCATTTGTATTAACCCGACTTGAAATTTGAGAGCAACAG 1806
 QY 581 AlaGlnLeuGlnGlnGlnLeuSerValLysAlaLysAsnLeuGlnMetValThrAsnSer 600
 DB 1807 CGTGAGAACTTTAAAAATTAAGAACTAGAACTAGAAATG---GAGCAA 1863
 QY 601 ArgGlnHisSerIleAsnAlaGlnValGln---ThrAspValGlnLysGlnValValArg 619
 DB 1864 AACCAAGAAATCAACAGAGGTGAATGGACGGATGCAGACAGAGGTGAGACCCACGC 1923
 QY 620 LysGlnMetSerValLeuGlnLysAspSerGlyTyrAsnAlaSerAsnSerAspLeuGlnAsp 639
 DB 1924 AAGGAGACGAATTCGG-----CAGCTCCAG--- 1950
 QY 640 SerSerValAspGlyLysArgLeuSerSerHisAspGlyCysIleGlnHisArgLys 659
 DB 1951 -----ATTGCAACGACGAGAGAGAGCCCAACGCCGAGC 1986
 QY 660 Met---LeuGlnGlnLysIleValAspLeuGlnGlnIleGlnLeuAsnLysLys 678
 DB 1987 TTCACATTCGAGAAACAAAGCTAAAGATTTA-----CTTGGCGAGAG 2028

QY 679 SerGluAsnAspLysGlnLysSerSerGlnGlnAsnPheMetGlnSerIleGlnLeuCys 698
 DB 2029 GAAAAATTTGAAGAGAGAGGCTGAGGAGACG----- 2061
 QY 699 GluAlaIleMetAlaGluLysAlaAsnAlaLeuGlnGlnLeuAlaLeuMetArgAspAsn 718
 DB 2062 -----CAGGAAATCGAGCTCCAAAGAAAGGA 2088
 QY 719 PheAspAsnIleIleLeuGlnAsnGluThrLeuLysArgGluIleAlaAspLeuGlnArg 738
 DB 2089 -----CAGGAAGAGAGCTTCTCCCGCCCAAGAAAGAACTCCACGA 2133
 QY 739 -----SerLeuLysGluAsnGlnGlnThrAsnGlnPheGluIleLeuGluLys----- 754
 DB 2134 CTCAAAGACTCAACACACAGAGAGGCTGAGATTTCAAGATTTTCAAGAACTGGAC 2193
 QY 755 GluThrGlnLysGlnHisGlnAlaGlnLeuIleHisGlnIleGlySerLeuLysLysLeu 774
 DB 2194 CAGCTCCAAAGAAAGAAATGACATGATGCC---AAGCTTGAACCTGAAAAAAGAGA 2250
 QY 775 ValGluAsnAlaGlu-----MetLysAsnGlnAsnLeuGlnLysLeu 790
 DB 2251 CTAGAGAGACGAGAGAGAGACAGCTCATGCTGTGCGCCATCTGGAAGACAGCTCCGA 2310
 QY 791 ThrLys-----ThrLysLeuLeuLysGlnGlnGlnIleGlnIleAlaGlnLeuArg 807
 DB 2311 GAGAAAGCAGAGATGATTCACACTCTCGCGCGTGGGAGGTACAGTGGGTGAAAGAGAG 2370
 QY 808 LysArgAlaAspAsnLeuGlnLysValArgAsnPheAspLeuSerValSerMetCly 827
 DB 2371 AAGAG----- 2376
 QY 828 AspSerGluLysLeuCysGlnGlnIlePheGlnLeuLysGlnSerLeuSer----- 844
 DB 2377 GACCTGGAAGCATTCGGGATCCCTCGGGGTGAAGAGGCTGTGCGGAGGGAT 2436
 QY 845 ---AspAlaGlnAlaValThrArgAspAlaGlnLysGlnCysSerPheLeuArgSerGlu 863
 DB 2437 GAAGATGGCGAGAGATTGAAGAAAGCTCACTGGCTTCTTCAATCAACAGAACGACAG 2496
 QY 864 AsnLeuGlnLeuLysGlnLysMetGlnAspThrSerAsnTyrAsnGlnLysGluLys 883
 DB 2497 CTTCCTCAAGCTAGCAACTTGGAGAGAGCTGTTCAGCAAGAAACATCTCAAAAA 2556
 QY 884 AlaAlaSerLeuPheGlnLysGlnLeuGlnThrGlnLysSerAsnTyrLysLysMetGlu 903
 DB 2557 GAAGTCCAAAGAAACAGAGAGATCTAGAGCTTAAATGTAACATGACAAAGAAATCT 2616
 QY 904 AlaAspLeuGlnLysGlnLeuGlnSerAlaPheAsnGlnIleAsnTyrLeuAsnGlnLys 923
 DB 2617 AGATTGTTGAAAAAATCATGATGAGAGTCTACAT----- 2652
 QY 924 LeuAlaGlyLysValProArgAspLeuSerArgValGlnLeuGlnLysLysValSer 943
 DB 2653 ---GTCACGAGAACTGCTCAAGATTTGAGAAATTAAGCCAGTGCAGTACAGCTCAA 2709
 QY 944 GluPheSerLysGlnGlnGlnLysAlaLeuGlnLysAsnAlaLeuGlnLysGlnVal 963
 DB 2710 TATAAAGACGCCAGCTACAG-----TACTCTCGCAAAATCACTTG 2751
 QY 964 ThrCysLeuSerGlu-----TyrLysPheLeuProAsnGlnValGluCys 978
 DB 2752 CCAACTGTGTTGGAAAAAGCAGAGACATTTTAAATTTCTTGACACAGGCCCTCTACG 2811
 QY 979 LeuLysAsn-----GlnIleSerLysAlaSerGlnGlnIleMetLeuLysGln 995
 DB 2812 TTAGACAAACACTCTTATCAAGTGAAGAAAGAAATGCAAGA----- 2853
 QY 996 GluGlnLysHisSerAlaSerIleIleSerLysGlnGlnIleIleMetGlnGlnSer 1015
 DB 2854 -----AAGAAACACAGCTTGCAACAGTACAGTACAGGCC 2883
 QY 1016 Glu-----GlnIleLeuGlnLeuThrAspGlnValThr-----His 1027

[illegible]

Oy	1048	AsplLeuPhagLysTyrTrlLeArgAsnLysSerGluAlaGlnAspLeuAArgGlnMet	1067
Db	2989	-----AGAAAGAAGCAGACGAGAGCGCCGTGACGGGCCCTG	3027
Oy	1068	GluAsnLeuLysGlyThrMetClnSerValGluValLysIleAlaAspThrLysHisGlu	1087
Db	3028	GCACGGCTGGAAGAGACATTCCTCGCTGCAG-----AGCACATCC	3069
Oy	1088	---LeuGlnGluThrIleArgAspLysGlnGluLeuLeuHisGlnLysTyPhePhe	1106
Db	3070	ACCCTGGCAGCAGCATTTGAGACGACGAGCG-----	3102
Oy	1107	GlnAlaMetClnThrIlePheProIleThrProLeuSerAspSerLeuProProSerLys	1126
Db	3103	-----AAA	3105
Oy	1127	LeuValGlnGlyAsnSerGlnAspProIleGluIleAsnAspTyrHisAsnLeuIleAla	1146
Db	3106	CATTGCCAGTCTGAACGTCGC-----	3126
Oy	1147	LeuAlaThrGluArgAsnAsnIleMetValCysLeuGlnGluThrGluArgAsnSerLys	1166
Db	3127	---AGCAGAGAGCAGTCAGAGCCTCCAGCGTAAGCTGAGAGCGACGACGAGAAAGCCCTGAG	3183
Oy	1167	GlnGlnValIleAspLeuAsnThrGlnLeuGlnInsSerGlnAlaGlnSerIleGluLys	1186
Db	3184	AAGGACCAAGCAGAGGTTAGAAATATGAATAATCCACAGCGCTGAACAAGATTTATAGAGTC	3243
Oy	1187	SerAspLeuGlnLysProLysGlnAspLeuGlnGlnGluVal---LysLeuLeuLeu	1205
Db	3244	GATGCTGTTAAAAAGATCATCATCGGAGGACCCGTGAAGAGAGTAGTGCTTCCAGCTTG	3303
Oy	1206	GlnMetGlnLeuLeuLysGlnLysIleThr-----AspSerGlnLeuSer---	1220
Db	3304	CCAAGTCAGTCGTGAAAAATCACCACCTGGTTCCTCCATGATGATGCCAGATCAATCCTTAC	3363
Oy	1221	IleGluLysLeuGlnLeuGlnAsnLeuGlnValThrGluLysLeuGlnThrLeu----	1238
Db	3364	ATTGAAAGAA-----GANGTCAAAGACGCTTCAGGATTTGCATGCT	3405
Oy	1239	-----GlnGlnGluMetLys	1243
Db	3406	GTGATTAGTGAAGCGTCAGTACATCGCACACAGATGAAGATTAATGAGAAACTTCAC	3465
Oy	1244	AsnIleThrIleGluArgAsnGlnLeuGlnInThrAsnPhelGlnLysLys-----	1260
Db	3466	AATGCACCAATTCACGTAACATAAATATAGAGCTGTGTGTACTCTCTGTCTCTCG	3525
Oy	1261	-----AlaGlnHisAspSerLeuLysGlnAspLeuSer	1271
Db	3526	ATGCCACAGCGCTGATGCCGCGCTCGGCGTAATCATCCCTTGCTCCAGCAAGATCTG--	3582
Oy	1272	GluAsnIleGlnInsSerIleGluThrGlnAspGlnLeu-----Arg	1285
Db	3583	----GTTACAGCTTCTCTTGATTTGGAAGAAACAGAAATCCCTGATTTAGTTTGCCAAAT	3636
Oy	1286	AlaAlaGlnGlnGluLeuArgGlnGlnLysGlnLeuValAspSerPheArgGlnGlnLeu	1305
Db	3637	GGAGTTCCAGGTGCATCCAAATTCACAGACATCCCTGGTTGACATGATT-----TACTT	3690
Oy	1306	LeuAspCysSerValGlyIleSerSerProAsnHisAspAlaValAlaAsnGlnGluLys	1325
Db	3691	CTTCATGGAATAATATGGAATGCAATGCTCCT-----	3720
Oy	1326	ValSerLeuGlnGluValAsnSerLeuGlnInsSerGlnMetLeuArgGlnGluArgAspGln	1345
Db	3721	---TCCCTGGCAGAACTTCAGTTACTGTCTTACACAAACAGTGAAGCATGGCGGTCT	3777
Oy	1346	LeuGlnThrSerCysLeuAlaLeuValSerGlnLeuGlnLeuLeuArgAlaHisValLys	1365
Db	3778	GGCATTGACCACTGCCAGTGCCTGATCTC-----CTTCGAAACCAACCATTTGCA	3825
Oy	1366	SerValGln-----	1368

Db 3826 CTGGGAGGAAGACTGTTTATCCAGCATTCGATCGAACAATACCTCTCCG 3885
 |||
 Qy 1369 GtGuanenleugluilethrllysleusnnglyleugluys-----Glu 1384
 |||
 Db 3886 GGTGCACATTTGATGATC--AAATGCCATGCTTAAAGTACAGTGTGTTT 3942
 |||
 Qy 1385 lleugluyssergluglusergluvalleuylsersermetleugluasleuysglu 1404
 |||
 Db 3943 GTTCCAGAAAAAAGTGTCAACAGTACAGTACTGCTTCAAGAACTCAAACT 4002
 |||
 Qy 1405 Aspsanlslyleuysglugluvalleuylserserlysgluasn-----Gln 1422
 |||
 Db 4003 TCAGTGGTTCAGAAATAGTCCACCTGACACCTTCAGAAAGCCCAATGTCAGTTG 4062
 |||
 Qy 1423 Pheserleugluvalleuylserserlysgluvalleuylsleuvalleuylleu 1442
 |||
 Db 4063 TTCCACACCCCATGATCTTCAAGGCGATGAGATGCGACCT-----GAGGTGG 4116
 |||
 Qy 1443 Lysalagluysalaleuylsleuylsleuylsleuylsleuylsleuylsleuyls 1453
 |||
 Db 4117 AAATCTTCAATCTCAAGATGAGCTCTT 4149

RESULT 10
 US-09-721-689-1
 : Sequence 1, Application US/09721689
 : Patent No. 6440685
 : GENERAL INFORMATION:
 : APPLICANT: Beraud, Christophe
 : APPLICANT: Freedman, Richard
 : TITLE OF INVENTION: No. 6440685el motor proteins and methods for
 : FILE REFERENCE: 1055
 : CURRENT FILING DATE: 2000-11-24
 : NUMBER OF SEQ ID NOS: 4
 : SOFTWARE: FastSeq for Windows Version 4.0
 : SEQ ID NO 1
 : LENGTH: 4176
 : TYPE: DNA
 : ORGANISM: Human
 US-09-721-689-1

Alignment Scores:
 Pred. No.: 1.65e-64 Length: 4176
 Score: 915.50 Matches: 416
 Percent Similarity: 40.59% Conservative: 246
 Best Local Similarity: 25.51% Mismatches: 536
 Query Match: 6.20% Indels: 433
 DB: Gaps: 71

US-09-150-867-1 (1-2954) x US-09-721-689-1 (1-4176)

Qy 6 AlavallysvalCyvalargvalargProleuileglnargglugluyspqlnala 25
 |||
 Db 7 TCAGTCCAGGAGGCGGAGGCGCGCCATGATCGACAGAAAGAGACTGAGAGCC 66
 |||
 Qy 26 Asn-----leuglntpllysalaglyAsnshhrlleserlval----- 39
 |||
 Db 67 AAGTTCATTATTCAGATGAGAGAAAGCAAAAGCAATCAAACTTAAACATACCGAA 126
 |||
 Qy 40 -----AspqlYthrlYserPheasnPheaspargvalPhe 51
 |||
 Db 127 GAGGACACTGGGAGCTCAGAGAAAGAACGACCAAGACCTACACTTATTTCTTT 186
 |||
 Qy 52 Asnserlsgluserlser-----GlnllyrYglgluile 64
 |||
 Db 187 TATCTCTGATACAAAAGCCAGATTCAGTTCACAGAAATGATTTCAAAACCTC 246
 |||
 Qy 65 Alavallprolelthargseralaleuylsleuylsleuylsleuylsleuylsleuyls 84
 |||
 Db 247 GGCACAGATGCTGAGAGTGTGATTTGAAGTTATATGCTGTGTGCTTGCATATGG 306

Qy 85 GlntrSerSerGlyserThrlPheMetGlyThrProAsnSerleuGlylleile 104
 |||
 Db 307 CAACATGATCTCGAAGACATACACTATGAGGGAATTCGAGATTCGCTTATA 366
 |||
 Qy 105 ProGlnAlaileGluValPheylsilleleGluile-----ProAsnarg 121
 |||
 Db 367 CCGGATCTGTAAGAGACTCTTCAGTCGATTAATGAAACACAGATGAGTACACT 426
 |||
 Qy 122 GlupheleleuArgValserYmetGluileYrAsnGluThrVallysaspheleu 141
 |||
 Db 427 TCTTTGCAATGAGTACGTACTTAGAAATTTAAACGAAGTGTGAGATCTACT 486
 |||
 Qy 142 CysAspAspArgArgLysLysPro-----LeuGluileArgGluAspPheasn 157
 |||
 Db 487 -----CGCGGAGAGTCATCTAAACCTTCATTTGAGTCCGTGAGATCCCAA 537
 |||
 Qy 158 ArgAsnValYrValAlaAspleuThrgluileuValMetValProGluHsValile 177
 |||
 Db 538 GAAAGCCCTTATGAGATTTATCCAAACATTTAGTACAGATTTATGTCAGCTAGA 597
 |||
 Qy 178 GlntrPleuysGlyGluysAsnArgHsYrGlyGluThrLysMetAsnshs 197
 |||
 Db 598 GAACTTATGATGCGGGCAATATCAACCGACCGACGCGATGGAAGACAGCTC 657
 |||
 Qy 198 SerSerArgSerHsThrlePhe-----ArgMetileValGluSerArgAspArgasn 215
 |||
 Db 658 ACTAGACAGTCTCATCCCATCTTCACATCAAGTTCACAGCTCAATTTATCTGAA 717
 |||
 Qy 216 AspProthrAsnSerGluAsnCysAspGlyAlaValMetValSerHsleuAsnleuVal 235
 |||
 Db 718 ATGCCA-----TGTGAA-----ACCGTCAATGATCACTGCTGTT 753
 |||
 Qy 236 AspleuAlaGlySerGluArgAlaSerGlnHrGlyAlaGluYrValArgleuysglu 255
 |||
 Db 754 GATCTGCGGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 813
 |||
 Qy 256 GlyCysAsnleuAsnArgSerleuPheleuGlyGluValleuylsleuSerasp 275
 |||
 Db 814 GGGGAAATATTAACAAAGCCCTTGTGACTGCGGGAAGCTATTCGCTTACTGAT 873
 |||
 Qy 276 GlyGlnAlaGlyly-----PheileAsnYrArg 285
 |||
 Db 874 TTATCTCAGATGCTCAAACTCTTCGAAAGAAAGCAAGTTCGCTGCTTACAGG 933
 |||
 Qy 286 AspSerlyleuThrArglleuGlnAsnserleuGlyAlaAsnAlaYrshrlvalile 305
 |||
 Db 934 GATTTGTTGACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 993
 |||
 Qy 306 lleCysThrleThrPro-----ValSerPheaspGluThrleuSerThrleuGlnPhe 323
 |||
 Db 994 ATTGCCACATTTCCACGCTGATGTCAATTAATGAGAAACCTTAAGTACTTGGCAT 1053
 |||
 Qy 324 AlaSerThrAlaLysHsValArgAsnThrProHsValAsnGluValleuAspAspqlu 343
 |||
 Db 1054 GCAATTAAGCCAAAACATCATCAACAGCTTCATTAATGAGATGCCAACCTCAAA 1113
 |||
 Qy 344 AlaileuysArgYrArgLysGluileuAspleuys----- 357
 |||
 Db 1114 ---CTTATCCGAGCTCGAGCTGAATAGCCAGACGTGAACGCTGCTGCTCAAGG 1170
 |||
 Qy 358 LysGlnleuGluAsnleuGluSerSerSerGluThrLysAlaGlnAlaMetAlaLysglu 377
 |||
 Db 1171 AATGAGATGGCCCTTATGACTCCGCCCA-----GCTTAAGTATGAG 1215
 |||
 Qy 378 GluHsThrGlnleu-----LeuAlaGluileLysGlnleuHsLysGluArgGluAspArg 396
 |||
 Db 1216 GAAAAACTTCAGCAAGATGAAGCAAGACTTCAGAAATGACCAAGATGACCAAAATAG 1275
 |||
 Qy 397 lleThrPheleuThrAsnilleValAlaSerSerGlnGluSerGlnAlaAspGlnArg 416
 |||
 Db 1276 ---TGAATGAAC-----CAAAAT 1293
 |||
 Qy 417 VallysArgLysArgArgValThrTrpAlaProGlyLyslleGlnAsnSerleuHsAla 436


```

Db 1294 ATTTGAAAGAACAACTCTA-----GCCCCAGAGAA 1326
Qy 437 SerGlyValSeraspPheaspMetLeuSerArgLeuProGlyaspNheSerIysLysAla 456
Db 1327 GAAGGGATTGGA-----GTTGTT 1344
Qy 457 LysPheSeraspMetProSerPheProGluIleaspServalysThr----- 473
Db 1345 TTGCATTTCTGAACTGCTTATTCATTTGATTCGATGATGACCTTTGAGTACTGGAAATC 1404
Qy 474 -----GluPheSeraspPheaspPalaLeuSerMet 484
Db 1405 ATCTTATATCATTTAAAGAGAGTCAGACATACGTTGCTGATGAGACATGCTTCCAGCGAG 1464
Qy 485 Metasp-----SeraspGlyIleaspAlaGlu-----TyraspLeu 496
Db 1465 CAAGATATTCCTTCCTTCGCTTTCGCTTTCGATTCGATTCCTTGGGAAAGACCAAT 1524
Qy 497 AlaSerIysValThrHisArgGlyLysThrSerLeuHisGlnSerMetIleaspPheGly 516
Db 1525 GGGGGGACAGTACT--CTGATACCCCTGAGTGGTCCAGTGTCTGTGAATGGTGT 1581
Qy 517 GlnIleSeraspServalGlnPheHisaspSerSer-----LysGluasp 531
Db 1582 CAGATCTGGAGGCGCACATCTAAATCAAGGTCCTGATTCCTTGGGAAAGACCAAT 1641
Qy 532 GlnLeuGlnIyr--LeuProLysaspSerGlyaspMetIleGlySargLys----- 548
Db 1642 ATGTTGGCTTTAACCTTCAACAAAGAGCCGCAAGCTCAGGAGAGAGAGAGAGTGGC 1701
Qy 549 -----AlaSerPheGlyLysGluIleThrSerLeuGlnGlnLeuGlnSerIysGlu 566
Db 1702 CTTCGTGCTCTCTTCACCTGCTGCATGACCGACCTC-----TCGAATCC 1746
Qy 567 GlnGluLysLysGluLeuValGln-----SerPheGluLeuLysIle 580
Db 1747 CGTGAGAAACCTGTCTGCATGCTGTATTAACCCGCGACTTGAATTTGAGAGCCAAACAG 1806
Qy 581 AlaGluLeuGlnGlnIleuSerValLysAlaLysaspLeuMetValThrAsnSer 600
Db 1807 CGTAAAGAACTTGAATAATTTGAAAAGTAAAGAACTCAAGAAAGAAATG--GAGGAA 1863
Qy 601 ArgGlnHisSerIleasnAlaGluValGln--ThraspValGlyGluValValArg 619
Db 1864 AAGCAGAAATCGACAGAGGCTGAACCTGAGCGATGCAGCAGAGGTGGAGAACCCAGCGC 1923
Qy 620 LysGluMetSerValLeuGlyaspSerGlyTyrAsnAlaSerAsnSeraspLeuGlnasp 639
Db 1924 AAGGAGACAGAAATCGTG-----CACTCCAG-- 1950
Qy 640 SerSerValaspGlyLysArgLeuSerSerSerHisaspGlyIleGlnHisArgLys 659
Db 1951 -----ATTCGCAAGCAGAGAGAGGACCTCAAAAGCGCGCAGC 1986
Qy 660 Met--LeuGlnGlnLysIleValaspLeuGlnGluPheIleGluaspLeuaspLysLys 678
Db 1987 TTCACATCGACAGAACACTTAAGGATTTA-----CTTGGCGAGAG 2028
Qy 679 SerGluaspaspLysGlnLysSerSerGlnLysaspPheMetGluSerIleGlnLeucys 698
Db 2029 GAAAAAATTTGAAGAGAGAGCTGAGGAGACAG----- 2061
Qy 699 GluAlaIleMetAlaGlnLysAlaasnAlaLeuGlnGluLeuAlaLeuMetArgAspAsn 718
Db 2062 -----CAGGAAATCGAGCTGCAGAGAGAGAGAG 2088
Qy 719 PheaspAsnIleIleLeuGlnaspGluThrIleuLysArgGluIleAlaaspLeuGlnArg 738
Db 2089 -----CAAGAGAGAGAGAGCTTTCTCCCTCCAGAGAGAACTCCACAGA 2133
Qy 739 -----SerLeuLysGluaspGlnGluThrAsnGluPheGluIleLeuGlnLys----- 754

```

```

Db 2134 CTCAAAGAACTCAACAGACAGAGAGCTGAGAACTTTCAGATATTTCAAGAACTGAGC 2193
Qy 755 GluThrGlnLysGlnHisGlnIleGlnLeuIleHisGlnIleGlySerLeuLysLysLeu 774
Db 2194 CAGCTCCAAAGAGAGAGAGAGATGACAGATGCTC--AAGCTTGAACCTGAGAGAGAGAGAG 2250
Qy 775 ValGluaspAlaGlu-----MetTyraspGlnaspLeuGlnaspLeuGlu 790
Db 2251 CTAGAGAGACAGAGAGAGAGAGAGAGTGCATGCTCGGCGCCATCTGAGAGAGAGCTCCGA 2310
Qy 791 ThrLys-----ThrLysLeuLeuLysGlnGlnIleGlnLeuAlaGluLeuArg 807
Db 2311 GAGAGACAGAGATGATTCAGCTCTGCGCGCTGGGAGAGTACATGGGTGAGAGAGAGAG 2370
Qy 808 LysArgAlaaspAsnLeuGlnLysValArgaspPheaspLeuSerValSerMetGly 827
Db 2371 AAGAG----- 2376
Qy 828 AspSerGlnLysLeuGlnGlnIlePheGlnLeuLysGlnSerLeuSer----- 844
Db 2377 GACCTGAAAGCATTCGGGAAATCCCTCGCGGGAGAGAGAGCTGCGGAGAGGAGAT 2436
Qy 845 -----AspAlaGluAlaValThrArgaspAlaGlnLysGlySerPheLeuArgSerGlu 863
Db 2437 GAGAGTGGCGAGAGATTAAGAAAGGCTCAACTGCTTTCTTGGAATTCAGAGAGAGCGAG 2496
Qy 864 AsnLeuGlnLeuLysGlnLysMetGluaspThrSerAsnTyrLysaspGlnLysGluLys 883
Db 2497 CTTCGCAAGTATGAACTTGAAGAGAGAGCTGTTCAACAGAAAGACATCTCGAAGAAA 2556
Qy 884 AlaAlaSerLeuPheGlnLysGlnLeuGlnIleThrLysSerAsnTyrLysMetGlu 903
Db 2557 GAAGTCAGAGAGAGAGAGAGATCTTGAAGTGTAAATGTGACATACAAAGAAATCT 2616
Qy 904 AlaaspLeuGlnIleGlnLeuGlnSerAlaPheaspGlnIleAsnTyrLeuaspGlnLeu 923
Db 2617 AGATTTGGAAGAAACATGATGAGCTGCACAGAT----- 2652
Qy 924 LeuAlaGlyValProArgaspLeuLeuSerArgValGlnLeuGlnLysValSer 943
Db 2653 ---GTCACGAGAACTGCTCAAGATTTTCAGAAATTAAGCCAGATGATACAGGCTGCAA 2709
Qy 944 GluPheSerIysGlnLeuGlnLysAlaLeuGlnLysAlaLeuGlnaspLysaspGlnVal 963
Db 2710 TATTAAGAAAGCCAGCTACAG-----TACTCTCGCAGAAATCACTTG 2751
Qy 964 ThrCysLeuSerGlu-----TyrLysPheLeuProaspGlnValGluLys 978
Db 2752 CCAACTCTGTTGAAGAAAGAGAGAGAGATTTGAATCTTGACAGAGGCTCTCAGC 2811
Qy 979 LeuLysAsn-----GlnIleSerLysAlaSerGlnGlnIleMetLeuLeuLysGln 995
Db 2812 TTAGACAAACACTTATATCAAGTAAAGAAATGAGAA----- 2853
Qy 996 GlnGlyGlnHisSerIleIleSerLysGlnGlnIleIleMetGlnGlnGlnIle 1015
Db 2854 -----AAGAGAAACAGCTTCACAGTACAGTACAGGCGC 2883
Qy 1016 Glu-----GlnIleLeuGlnLeuThrAspGluValThr-----His 1027
Db 2884 AATGCAAAACAGCTCAAAAGCTCAAGCCACTTGAATTCATCTGCACACATTCGACGT 2943
Qy 1028 ThrGlnSerLysValGlnGlnThrGlnGlnGlnIleTyrLeuGlnMetLysLysMetHisasp 1047
Db 2944 CAGAGAGAGAGAGAGAGAGAGAGAGAGATTTTGGAGTCC----- 2988
Qy 1048 AspLeuPheGlnLysTyrIleArgaspLysSerGlnAlaGlnaspLeuLeuArgGluMet 1067
Db 2989 -----AAGAGAAACAGCAGAGAGAGAGCGCTGAGAGCGGCGCTG 3027
Qy 1068 GluaspLeuLysGlyThrMetGluSerValGluValLysIleAlaaspThrLysHisGlu 1087
Db 3028 GCCAGGCTGAGAGAGAGACATTCGCGCTGCAG-----AGGCACTCC 3069

```

OY	1088	--LeuGIuGIuThrIleAArgAspLysGIuGIuLeuHisGIuLysTyrPhePhe	1106
Db	3070	ACCCGTGGGCACGAGATTTGAAGACGAGAGCAG	3102
OY	1107	GlnAlaMetGIuThrIlePheProIleThrProLeuSerAspSerLeuProProSerLys	1126
Db	3103	-----AAA	3105
OY	1127	LeuValGIuGIuAsnSerGlnAspProIleGIuIleAsnAspTyrHisAsnIleAla	1146
Db	3106	CTTCCAGCTGTGAACAGTGC	3126
OY	1147	LeuAlaThrGIuArgAsnAsnIleMetValLysLeuGIuThrGlnArgAsnSerLeuLys	1166
Db	3127	---AGCAAGACGAGCTGAGGCGTCCAGAGCTAGCCTGAGAGCTGACACAGAGAGAGCCCTTGAG	3183
OY	1167	GIuGIuValIleAspLeuAsnThrGlnLeuGIuSerLeuGlnAlaGlnSerIleGIuLys	1186
Db	3184	AAGGACGAGGAGAGGTAGAAATATGAATATCGACAGACTGAACAGAGATTATAGAGTC	3243
OY	1187	SerAspLeuGIuLysProLysGlnAspLeuGIuGIuGIuVal---LysLeuLeu	1205
Db	3244	GATGCTGTCAAAAAGATCATCATGTGGACCTGGAAGGAGAGAGTGCTTCCAGCTTG	3303
OY	1206	GIuMetGIuLeuLeuLysGIuHisLeuThr-----AspSerGlnLeuSer-----	1220
Db	3304	CCAGTCAGTGTCTGAAAATATCACACGTCGTTCCTCCATCATGATGACGAGATCAATGCTTAC	3353
OY	1221	IleGIuLysLeuGIuLeuGIuAsnLeuGIuValThrGlnLysLeuGIuThrLeu-----	1238
Db	3364	ATTGAAAGAA-----GAAATCCAAAGACGCCCTTCAAGATTTTGATGCT	3405
OY	1239	-----GlnGIuGIuMetLys	1243
Db	3406	GTGATTAGTGAAGCTGTCACTACATCTGCAGACAGCATAGAGATATAGAGAACTTAC	3465
OY	1244	AsnIleThrIleGIuArgAsnGIuLeuGIuThrAsnPheGIuAspLeuLys-----	1260
Db	3466	AATGGCACCTTCAACGTAACATACTAAATATGAGCTGTGCTGACCTCTGTGTCTCTG	3525
OY	1261	-----AlaGlnHisAspSerLeuLysGlnAspLeuSer	1271
Db	3526	ATGCCAGACCTGATGCCCTGCTGCTGCTGATATCATCTCTTCCACGACGAGATCTG--	3582
OY	1272	GlnAsnIleGlnIleGlnSerIleGlnThrGlnAspGIuLeu-----Arg	1285
Db	3583	-----GTTCAAGCTTCTCTGATTTGGAAACAGAAATCCCGATTAGTTTGCCAAAT	3636
OY	1286	AlaAlaGlnGIuLeuLeuArgGIuGIuLysGlnLeuValAspSerPheArgGlnGlnLeu	1305
Db	3637	GGAATTGAGGTGATCATCCAAATTCGAGACTACCTTGCTTGACATGAT-----TACTTT	3690
OY	1306	LeuAspLysSerValGIuLysIleSerSerProAsnHisAspAlaValAlaAsnGlnGIuLys	1325
Db	3691	CTTATGTGAATATGGAAGCATGTCCCT-----	3720
OY	1326	ValSerLeuGIuLysValAlaAsnSerLeuGIuSerGlnMetLeuArgGIuLysArgAspGIu	1345
Db	3721	---TCCCTGCGACAGAAAGTTCAGTTACTGCTCTACACAAACAGTAAGATGATGGGAGCTCT	3777
OY	1346	LeuGIuThrSerLysValAlaLeuValSerGIuLeuGIuLeuLeuArgAlaHisValLys	1365
Db	3778	GGCCATGACCACTGCGAGTGGCTAGTC-----CTTGTAAACCCACCATTTGCA	3825
OY	1366	SerValGIu-----	1368
Db	3826	CTGGTGAAGAGACACTGTGTTTTTATCCACCATTCGATCTCGAAATACATCTCTCCG	3885
OY	1369	GlyLysAsnLeuGIuLysIleThrLysLysLeuAsnGlyLeuGIuLys-----Glu	1384
Db	3886	GTCTCACAAATTTATGTGATC---AATGCACTGCTTTAATGAAATTCAGGTGTGTGTT	3942

```

QY 1385 IIELEUgLyISserGIuISerGIuValLEuISerNetLEuGluAnLEuLyGIu 140
      ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3943 GTTCCAGAAATAAATAATGTGTCAACACTAGTCTTCTTACAGAAACTCAACCT 400
      ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1405 ASpanAsnLyLEuLySGIUGInAlGIuGIuLyISerSerLySGIuAn-----GIn 142
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4003 TCAGTGGGTCCAGAAATATGTCCACCTTGACGACCTTACGAGAGCCCAAAATGTCCACTTG 406
      ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1423 PheserLEuGIuValPheserGIuISerGIuISerValAspGIuIIEGIuValLEu 144
      ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4063 TTCACACCCCCCATGTATCTTCAAGGCGATCAGATGTCCACCT-----GAGGTGTGG 411
      ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1443 LysAlaGIuLEuLyAlaAlaGIuGIuLyGIu 1453
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4117 AAACCTTACTTCATTCATGATGAGGTGCTCT 4149
      ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
US-09-592-054-1
; Sequence 1, Application US/09592054
; Patent No. 6440684
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Finer, Jeffrey
; APPLICANT: Sakowicz, Roman
; APPLICANT: Wood, Kenneth
; TITLE OF INVENTION: No. 6440684el motor proteins and methods for
; FILE REFERENCE: 1016
; CURRENT APPLICATION NUMBER: US/09/592,054
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 4308
; TYPE: DNA
; ORGANISM: Human
US-09-592-054-1

Alignment Scores:
Pred. No.: 5,3e-64 Length: 4308
Score: 909.50 Matches: 368
Percent Similarity: 42.44% Conservative: 275
Best Local Similarity: 24.29% Mismatches: 503
Query Match: 6.16% Indels: 369
DB: 4 Gaps: 61

US-09-150-867-1 (1-2954) x US-09-592-054-1 (1-4308)
QY 3 GIuGIyAspAlaValLyLyValCyValArGIuProIIEGIuArGIu----- 20
      ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 76 AAGGAAATTCCTGTAAAGAGTGGCGCTGTGCGCTGTGCGCTGTGCGCCCAAGAGATTAGC 135
      ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 21 GIuGIyAspGIuAlaAsnLEuGIuTriPLyAlaGIuAsnAnThrIIEserGIuValAsp 40
      ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 136 GAGGGCTGCCAGANGTGGCTTCTCTGTGCGCCGAGAGGCTCAGTG-----GTGGTT 189
      ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 41 GIyThr---LysSerPhesAnPhesAspArGIuPhesAnSerHisGIuSerThrSerGIu 59
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 190 GGTATCAGATTAATATCTCCACCTACGATTTGTATGTATGATCCCTCTACTGAAACAGAAAGAA 249
      ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 60 IIEyGIuGIuIIEAlaValProIIELEySerAlaLEuGIuGIuLyTrAsnGIuLyThr 79
      ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 250 GCTTCATATACAGCAGTAGGCCCACTCATTAAGGTATATTAAAGGATTAATATGCAACG 309
      ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 80 IIEpheAlaTrGIuGIuThrSerSerGIuLySThrTrThrNetwEgLy----- 96
      ::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 310 GTCCGTGCGCTATGGGCGACAGACTGCTGTGAAATAACATTCATTCAGGAGGTGCATATACT 369
      ::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 97 -----ThrProAnSerLEuGIuLEIIEProGInAlIIEGIuGIuValPhe 112
      ::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 370 GCAGAGCAAGAAATGAACCAAGTGGGGTATTTCTTGGAGTAAATACAACTGCTCTTC 429
      ::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 113 LysIIEIIEGIuGIuLEIIEProAnArGIuPhuPheuLEuArGIuIIEserTrIIErMetGIuIIE 1322

```



```

Db 1700 GTACCCAGAGCTCAAGAGAGAGTGTAGAG-----CAGAGGAAGTCAAT 1747
QY 543 TAlAGlucysArgLys-----AlaSerPheGluLysGluLeuThrSerLeuGluGlnG 561
Db 1748 CGCTACCCCTGGGCTGAGTGTAGAGCTTTGCACTGGAGGCGCCGAAATCCACAGCA 1807
QY 561 nLeuGlnSerLysGluGluGlnLysLysGluLeuValGlnSerPheGluLeuLysIleAl 581
Db 1808 GAGCAGATCCCGCAAGAAAGAGGTAAAGAGAGTGTCCAGAGCTCCGAA----- 1856
QY 561 aGluLeuGluGluGlnLeuSerValLysAlaLysAsnLeuGluMetValThrAsnSerAr 601
Db 1857 -GAGCTGACTGTAAACTACGACCAAGAAATCCAGAGAGTGTAGT----- 1898
QY 601 gGlnHisSerIleAsnAlaGluValGlnThrAspValGluLysGluValArgLysG 621
Db 1899 -----AACAGAAACAGAGATATGATGCTCCACAGAGAGAGCTGCACACAGAA 1945
QY 621 uMetSerValLeuGluLysAspSerGlyTyrAsnAlaSerAsnSerAspLeuGlnAsp 639
Db 1946 GCAGCTGTGT-----TTCCAGCGCGCTCCACAGAGTGTACAGACACTCAA 1990
QY 640 -----SerSerValaSerGlyLysArgLysSerSerSerHisAspGluLysIleGlnH 657
Db 1991 TGACATGCTCTCACACACAGAAAGAGCGATCAG----- 2024
QY 657 sArgLysMetLeuGluGlnLysIleValaLeuGluGlnPheLeuGlnAsnLeuAsnLys 677
Db 2025 -----GAAATGCTAACCAACCTACTGCGGAGCTCGGCCAAGTGGGCCAGGCCATTGCC 2080
QY 677 sLysSerGluAsnAspLysGlnLysSerSerGluGlnAspPheMetGluSerIleGlnLe 697
Db 2081 CGCGGAGTCCAGCATCGACCTTAAGATGAGT----- 2111
QY 697 uCysGluAlaIleMetAlaGluLysAlaAsnAlaLeuGluGln----- 711
Db 2112 -----GCTGTGCTGCGACAGGATGCCAGAGGTGAGAGAAATTTACACCATGGCCG 2164
QY 712 -----LeuAlaLeuMetArgAspAsnPheAspAsnIleLeuGlnAsnGluThrLe 729
Db 2165 TTTGTTATTCAGCAAGATGAAGAGCGAGGCCAAGACATTCGCCAGCATGCTCCACAT 2224
QY 729 uLysArgLysIleAlaAspLeuGluArgSerLeuLysGlu-----AsnGlnGluThrAsnG 748
Db 2225 GGAACACAGACAGGCTGACTCCACACAGAGATCTCCGAATATGAGAAAGATCTGGCGGA 2284
QY 748 uPheGluIleLeuGluLysGluThrGlnLysGlnHisGluAlaGlnLeuLeuHisGluI 768
Db 2285 GTACCGGCTACTC-----ATTTCGACGACGAGCGACGCGCATG----- 2321
QY 768 eGlySerLeuLysLysLeuValGluAsnAlaGluMetLysAsnGlnAsnLeuGluGlnAs 788
Db 2322 -AATTCCTGTCAGAGATGATGCGGAGCGAGAGACAGAAAGAGCGACCTCGAGAGACA 2380
QY 788 pLeuGlnThr-----LysThrLysLeuLeuLysGluGlnGlnIleG 802
Db 2381 AATGATTCGCTGCGCGAGAAATGCGCCAAAGCTCAAGGCCCGCGAGCAGCTTCGCGCT 2440
QY 802 nLeuAlaGluLeuArgLysArgAlaAspAsnLeuGlnLysLysValArgAsnPheAspLe 822
Db 2441 TAAAGCCGAGAGAAAGCGGCTGAGAGACTCGC----- 2477
QY 822 uSerValSerMetGlyAspSerGluLysLeuGluGlnLeuPheGlnLeuLysGlnSe 842
Db 2478 -----TCCATGTTTCATTTTCAG-----ATGAGACAGCTA----- 2507
QY 842 rLeuSerAspAlaGluAlaValThrArgAspAlaGlnLysGluLysSerPheLeuArgSe 862
Db 2508 -----CGCGAAGGCCACACCGG-----CAGGTGTCGAGCTCCGGGA 2545
QY 862 rGluAsnLeuGluLysGluLysMetGluAspThrSerAsnTrpLysAsnGlnLysG 882

```

```

Db 2546 CGAAATTCGCCGCAAGCAGCAGCAAGAAATGACAGATGAAGAT---GTCCATCAAAAG-- 2600
QY 882 uLysAlaIleSerLeuPheGluLysGlnLeuGluThrGlnLysSerAsnTrpLysLysMe 902
Db 2601 -----CTGCTCTTGGCGCACCAACAGATGACGGCCGCTACGAAAGGTGCGCCAGA 2653
QY 902 tGlnAlaAspLeuGlnLysGluLeuGlnSerAlaPheAsnGlnIleAsnTrpLeuAsnG 922
Db 2654 GGATCCCGAAGATCCACAGAGCTTCAG-----AACAT 2686
QY 922 yLeuLeuAlaGlyLysValProArgAspLeuLeuSerArgValGluLeuGluLysVal 942
Db 2687 CATCTTCAC----- 2696
QY 942 lSerGluPheSerLysGlnLeuGluLysAlaLeuGluGlnLysAsnAlaLeuGluAsnG 962
Db 2697 -----AACAGCTCTCGGAGCAAGCGCGCAAAAGCTCAAGAGGCTCGAGGAC-- 2744
QY 962 uValThrCysLeuSerGluTrpLysPheLeuProAsnGlnValGluLysLeuLysAsnG 982
Db 2745 -----ACGGTGCCCAAGAGATGTCAGACCTACACAACTCGCAAAACTTTGCTCAGA 2800
QY 982 nLysSerLysAlaSerGluGluIleMetLeuLeuLysGlnGluGlnLysIleHisSerAlaSe 1002
Db 2801 TCTACAGCAACGATCCGAAGAAATGCTTAACGAGAGAGACGAGAGAGCGGTGATC 2860
QY 1002 rIleLeSerLysGlnGluIleIleMetGlnGluGlnSerGluGlnIleLeuGlnLeuTh 1022
Db 2861 ACTCGCGCAAGAAACAGAAATTTCTTCTTGAGAACAAAC-----CT 2902
QY 1022 rAspGluValThrHisThrGlnSerLysValGlnGlnThrGluGlnLysIleGluLys 1042
Db 2903 CGACAGCTGAC----- 2915
QY 1042 tLysLysMetHisAspAspLeuPheGluLysTrpIleArgAsnLysSerGluAlaGlnAs 1062
Db 2916 ---AAGGTGCAACAGCAATG-----GTGGGGAACAGCGCGATCTCGCGTG 2959
QY 1062 pLeuLeuArgGluMetGlu-----AsnLeuLysGlyThrMetGluSerValGluValLysI 1081
Db 2960 CGAGCTGCCCAAGCTGGAGAAAGCGTCTACGCTTACCTGATGAGCGGTGAAGGCTCTGGA 3019
QY 1081 eAlaAspThrLysHisGluLeuGluGluThrIleArgAspLysGluGlnLeuLeuHisG 1101
Db 3020 GACAGCGCTCAAGAGAGCGGAGAGAGCGCATGCGGATGCGCAAGCGCTACCAATAGCA 3079
QY 1101 u 1101
Db 3080 G 3080

RESULT 13
US-09-134-001C-322
Sequence 322, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT FILING DATE: US/09/134,001C
CURRENT APPLICATION NUMBER: 1998-08-13
PRIOR FILING DATE: US 60/064,964
PRIOR APPLICATION NUMBER: 1997-11-08
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 322
LENGTH: 30549
TYPE: DNA
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-322

Alignment Scores:

```

Pred. No.:	3,265-59	Length:	3054.9
Score:	866.00	Matches:	661
Percent Similarity:	38.24%	Conservative:	697
Best Local Similarity:	18.61%	Mismatch:	1295
Query Match:	5.86%	Indels:	898
DB:	4	Gaps:	157

US-09-150-867-1 (1-2954) x US-09-134-001C-322 (1-30549)	
QY 15	ProLeuIleGlnArgGluGlnGlySerGlnAlaIAsnLeuGlnTrpLys-----30
DB 16867	CCMAATCAAGAACMAAAATGAATTAATGAGCAGCGCTAATTAATCAAAATCAATTAATGACACTCAAAACA 169236
QY 31	-----AlaGlyAsnAsn-----ThcIleSerGlnValAsp 40
DB 16927	GCGTTAAAGTGGAGAAATTAATTAATTAACACACTGACACCAAGAAAGCACAATAATTAATGACAAATGACAA 169866
QY 41	GlyThrLysSerPheAsnPheAsnArgVal-----PheAsnSerHis 54
DB 16987	GSTTATCTAGCTTGAACACAGCCTCAATTAACAGCGCGCAAAAAGATTATAGTAATCAATCAAGCT 170466
QY 55	GluSerThrSerGlnIleTyrGlnGluIle---AlaValProIleIleArgSerAlaLeu 73
DB 17047	AAAAACAAGACGATGGTGTGCTCAAAAGTAAAGTAAAGTAAAGTAAAGTAAATTCCTCTATG 171066
QY 74	GlnGlyTyrAsnGlyThrIlePheAlaTyrGlnThrSerSerGlyThrTyrThr 93
DB 17107	AGTAATTAAAGAGATGGCAATTAATAAAGAGACATCAACGTAATGACGTCATAT--- 171633
QY 94	MetMetGlyThrProAsnSerLeuGlyIleIleProGlnAlaIleGlnGluValPheLys 113
DB 17164	ATCAACGACGATCCGACATTAAGTTACAGCTTACGATTAACAGCACTACAGAACGACAGAAAT 172233
QY 114	IleIleGlnGluIleProAsnArgLysPhe-----LeuLeu 125
DB 17224	ATCAATCAATGATCCACACCAACGTAAGGCTTAATAAGCTACAAATGAACACAGCGCTATCA 172833
QY 126	ArgVal-----SerTyrMetGluIleTyrAsn-----134
DB 17284	GCGGTTCACAAACACACAAACAAAGATGCTGTATGGTGTTCACAAATTAATGCTAATGCTAAACA 173433
QY 135	-----GluThrValIleLysAsnLeu---LeuCysAspAspArgArgLysProLeu 150
DB 17344	CAAGCTACACAAACCTCATTAATGGTTAAATTAAGCTTAATAAAGACGGTCAAAAGCGTGAATTA 174033
QY 151	-----GluIleArgGluAspPheAsnArgAsn 159
DB 17404	AATCTATTATTAATTAATCACTAATACCCGCTACAAAAGATCAACAAAGAAATTAACAAAGCA 174633
QY 160	ValTyrValAlaAspPheThrGlnGluLeuValMetValProGlnHisValIleGln---178
DB 17464	ACTGATATGAACACGCCGATGGAAGGTTTAAAGAAACAGTTCACAAAACCTGTGATCAAGTA 175233
QY 179	-----TrpIleLysGlyGlyLysAsnArgHisTyrGlyGlnThrLys 193
DB 17524	AAACAAGTAGCAATTAATGTCTCATGATGATCAACTGCAACACACACATTAATGATATGCT 175833
QY 194	MetAsnAsp-----HisSerSerArgSerHisThrIlePheArgMetIle 208
DB 17584	GTCATATGAAGCTCAACGCTCAATCAACACAAATGCTCAACCTGTTCTTAACAAATTAAGCT 176433
QY 209	ValGluSerArgAspArgAsnAspProThrAsnSerGlnAsnCysAspGlyAlaValMet 228
DB 17644	ATAGAACGTTTAACCTCAACCTGTAAACACTACAAAAGATGCAATTAATGCTGCT----- 176933
QY 229	ValSerHisLeuAsnLeuValAspLeuAlaGlySerGluArgAlaSerGlnThrGlyAla 248
DB 17698	-----CAAAAAGCTGACACAAAGACCAACAGCTGCTGAACCTGGAATA 177393
QY 249	GlnGlyValArg-----252
DB 17740	GCGTGTAAACAGTCTCAATGACAACTCAGAAAATGCTGAAGTAGCTTAAGTAACTGCA 177993

QY	253	-----	Leuysgluylcysasnileasntrser	262
Db	17800	GCAACAACAGCGATGAAAGTGGAAATTTGTGTCAAAGGAAGCAACAATGTAGATCTGCA	17859	
QY	263	Leupheileleuglyglvalillelyslysleuseraspglygl--Alaglyglyphe	281	
Db	17860	ATGCTTGTTTACGTAAAAAGCAATTAAGATATAAAACGTAATAATAGTAGTAAAT	17919	
QY	282	ileasnlytargdaspserselysleuthrarglileuglnasnserleuglyglasna	301	
Db	17920	ATTAAAT---GAGGATCATGACCAACAACAGCTTATGACATTCGCTGA---AATTAATGCT	17973	
QY	302	lysthrvalillelecsythrllethrprovalserpheaspgliuthrleuserthleu	321	
Db	17974	CAACAAGTTATC-----GATGAACCTCAAGCAACGTTA	18006	
QY	322	-----glphealaserthrlyalylshisvalargdsnthrprohs	335	
Db	18007	ACGTACAGATACAAATCAATTCATATGGCAAAATGCCGTAACTCAAGCTAAATCTATCTTCAT	18066	
QY	336	valaenglivalileuaspaspargliulaleuylsargtytarglysgliuleuasp	355	
Db	18067	GGAGATACATAACTCAACACAGAT-----AAAGATAGCTTAACCAACAGAT---	18114	
QY	356	leuylslysglnleuglnasleucluserserserglythrllyalaglnalmetala	375	
Db	18115	-----GCTCAATTCACAGAAATTTGAATTCAGCTCAAAAACATATGCAAGATTTCTTAAT	18168	
QY	376	lysglyglunhsthrgln-----	381	
Db	18169	GATATATGATCTACACGTACGCAAGTCCAAACACAGATTTAACAGAGCTCAACCTTAGAT	18228	
QY	382	---leuleualagluillelysglneuthslysgluarglyuasprgltlethrhisleu	400	
Db	18229	GGTTTAATGGTGGCCCTTAAMAGAAATATTAAGAT-----TAT	18267	
QY	401	thrpsnilevalval-----	405	
Db	18268	ACTAATATTTGTTCAACGCGTAATTACATCAATGGGGAACGCTCTACAGAACACATAT	18327	
QY	406	---AlaserSerglngluserglngl-----Aspglnargvallysarg	419	
Db	18328	GATGACAGCTGTCAAAAATGCTCAAAATATTAATTAATGCAACGCAACCAACAATTAAT	18387	
QY	420	lysargargvalthrthrleuprogluslyleglnasnserleuhsila-----	436	
Db	18388	AAAGGTATGTCTACACACACACAAACCGTGAATAATACTAAAGATGCTTAAACGGT	18447	
QY	437	-----Serglyvalseraspsrphemeluserarg	447	
Db	18448	GATCATAGATTAGAGGAACGTAAATAATGCGCAATCAAAACAATCGAATCTATCTAAT	18507	
QY	448	leuproglyasnpheserlyslalalys-----pheseraspmetroserphe	464	
Db	18508	TTG---AACAAATGCCCAAAAMATCGACAGGAAAAATCTAGTTAAAGCGCATCAACATTA	18564	
QY	465	progluileaspraspervalcysythrgluipheseraspheaspralaleuser---	483	
Db	18565	GAAACAATTCAACAAATCTTACAACCGCT---CAACAATTGATTAATGCTATGGGTGAG	18621	
QY	483	-----	483	
Db	18622	TTACGCAAAAGATTTGCTAAAAAGATCAAGTGAAGCAAGATAGTAAATATCTAAATGAA	18681	
QY	484	-----MetmetaspsersasnlylleaspralagluTPrasnleuAlaserlysalthr	501	
Db	18682	GATCTCAAAATTAAGCAAAATCATGATGATGCGAGTTCACACGTGTGAACCTTTATTAAAC	18741	
QY	502	hisargglyluthserleuhsglnsermettleasprheglyglilleaserpspsr	521	
Db	18742	GAAACTCAAAACCCGTGAATTCCTTAAGCAACATTTGAC-----CAAGCAACTCAATCC	18795	
QY	522	valgln-----Phehisaspserselysgluasnlgln---leuqlnlyr	535	

QY 1098 LeuLeuHisGluIuYsLysTyrPhe----- 1105
 Db 20914 ACCTAAATCTCTAGTACTTCTCAATGAGATTCGAGAAAAAGTTAGCGTACGATAT 20973
 QY 1106 -----PheGlnAlaMetGlnThrIlePheProIleThrProLeuSerAspSerLeuPro 1123
 Db 20974 GCTGTAGCCAGCTGAGCAACTC-----ATAATCAACTTAACGACCCCAACTATG 21024
 QY 1124 ProSerLysLeuValGluGluYasnSerGlnAspProIleGluIleAsnAsp---TyrHis 1142
 Db 21025 GATRTAAGTAAATCTCAAGCTATTCTCAAAAGGTCATTCAGCAAAAGATTCATTGCAC 21084
 QY 1143 AsnLeuIleAlaLeuAlaThrGluArg-----AsnAsnIleMetValCysLeuGluThr 1160
 Db 21085 GGTGCGAATAAGTCTGCACAAAATCAAGCAGATTCAAATTTAATAATCAATCACA 21144
 QY 1161 GluArgAsnSerLeuYsGluGlnValIleAsp----- 1171
 Db 21145 AATTTAATGATTAACAAAGCAAGCATTTAATGACTTAATATCATGCTCAAACTTAA 21204
 QY 1172 -----LeuAsnThrGlnLeuGlnSerLeu 1179
 Db 21205 CAGCAAGTGGCAGAAATATTCGACAAAGCTAATAGTTAATTAACGAATGGCGACTA 21264
 QY 1180 GlnAlaGlnSerIleGluYsSerAspLeu---GlnYsProIysGlnAspLeuGlu 1198
 Db 21265 AAAACACTGCTAGACAGACAGTCAACGTTTCATCAACAAAGTAAATATTAATGAGAT 21324
 QY 1199 GlyIuValIysLeuLeuLeuGluMetGlnLeuLeuYsGlnHis----- 1213
 Db 21325 CCGCAAGTCTAAATATTTATTAATGACTCCATCAAAAAGTCGAGAAATATTAAACGGC 21384
 QY 1214 LeuThrAspSerGlnLeuSerIleGluYsLeu-----GlnLeuGluAsnLeuGluVal 1231
 Db 21385 ACACAGATGATGTTTAAACAAACAAATTAACAGATGCGATTCGCAAAATTCATTTA 21444
 QY 1232 Thr-----GluYsLeuGlnThrLeuGlnGluMetLysAsn 1244
 Db 21445 ACATAAAACGATTTCATGCTGATCAAAAATTCAAAACCAACAGATGCAACCAAT 21504
 QY 1245 IleThrIleGluArgAsnGluLeu---GlnThrAsnIleGlnAspLeuYsAlaGluHis 1263
 Db 21505 -----GAAATTAACCTATTTAAACAAATCTAAACAAATTCACAAAGCAGGACAT 21555
 QY 1264 AspSerLeu-----LysGlnAspLeuSerGluAsnIleGluGlnSerIle 1278
 Db 21556 GATGAGATTAACCTGCTGCTCTCAAGAACGAGCTTCTAATGATTTAAATCATGCTAAA 21615
 QY 1279 GluThrGlnAspGluLeuArgAlaIleGlnGluLeu-----ArgGluGlnLys 1295
 Db 21616 GCACTTAATGAGCTATGCTCAACTGAGATGAATGAGTGTCTTGAACACAGTGTAAA 21675
 QY 1296 GlnLeuValAspSerPheArg----- 1302
 Db 21676 AAATTAAGCGACTTATCAATGAGATGAAGCGGACAAAATGAAATATGATGACATT 21735
 QY 1303 GlnGlnLeuLeuAspCysSerValGlyIleSerSerProAsnHisAspAlaValAlaAsn 1322
 Db 21736 CAAAAAGCTAAAGACATTAACACGGGCTCCAGTAGTACGCTTAAGTTAAAGCTACAAAT 21795
 QY 1323 GlnGlnLysValSerLeuGluValAsnSerLeuGlnSerGluMetLeuArgGlyGlu 1342
 Db 21796 GAAGAT-----GCTTATATGAAATGCAAAATGCTAGA---GAAAGTTTACATGTGTAG 21846
 QY 1343 ArgAspGluLeuGlnThrAspCysLysAlaLeuValSerGluLeuGluLeuValArgAla 1362
 Db 21847 CAA---AACTGTAAGAGGCTAAATAATCAAGCTGTGCTGAAATGATTAATTAACAAGCA 21903
 QY 1363 HisValLysSerValGluGluYasnLeuGluIleThrLysLysLeuAsnGly----- 1380
 Db 21904 -----TTAAATCTCTGACAGC 21918

QY 1381 -----LeuGluYsGlnIleLeuGluYsSerGluIuSerGluValLeuYsSerMet 1398
 Db 21919 GTTCTTCTGAAAAAACAATAGTTAAACCAAGCATTCACCAACCAAGAAAGTTCAAGAACCC 21978
 QY 1399 LeuGluAsnLeuYsGluAspAsnAsnLysLeuYsGluGlnAlaGluIuTyrSerSer 1418
 Db 21979 TTACAAAAGCAAAAGAACTTAATGAAGCTATGAAACACTGAAAACTGAAATAATTA 22038
 QY 1419 LysGluAsnGlnPheSerLeuGluGluValPhe-----SerLysSerGlnLys 1434
 Db 22039 AAGAA---CAATCAAGCTGATAGTATGATTAATGCTGACAGCTGGCTTCAAGCA 22095
 QY 1435 LeuValAspGluIleGluValLeuYsAlaGlnLeuYsAlaAlaGluIuArgLeuGlu 1454
 Db 22096 AATTCAATTCGCTGATTAATTAATGCTTCAATTAATTCGCAACTCCCAACCAAGCAG 22155
 QY 1455 IleLysAspArgAspTyrPheGluLeuValGlnThrAlaAsnThrAsnLeuValGluGlu 1474
 Db 22156 CTTAATTAAGATGTAATTAATTAAGCAACTCAACGATTTAAACT----- 22200
 QY 1475 LysLeuGluThrProLeuGlnAlaAspHisGluGluYsSerIleAspArgSerGlu 1494
 Db 22201 -----GCTGAAATATTTAAATGGGCAATCTTAA 22230
 QY 1495 GluMetGluIleYsValLeuGlu---GluYsLeuGluArgAsnGlnTyrLeuLeuGlu 1513
 Db 22231 TTAGCAGAGGCTAATGACAGCGAAATCAACAGCATGCAACATTTGCAAGATTAACCA 22290
 QY 1514 ArgLeuGlnGluIuYsLysLeuGluLeuSerAsnLysLeuGluIleLeuGln----- 1530
 Db 22291 TCACAAAAGATTAACCAACATGATTAATTAATCAAGCTCAAACTAAACCAAGATGAT 22350
 QY 1531 -----LysGluMetGluThrSerValLeuLeuYsAspArgLeuGln 1544
 Db 22351 GATATGCTAATTAATCTTAACCAATAGATTAATCTATG-----ATCAACATACAA 22401
 QY 1545 GlnYsLeuGluSerLeuLeuSerGluAsnIleIleLeuYsGluAsnIleAspThrThr 1564
 Db 22402 CAATGCTT-----AACATGCAATTAACA 22425
 QY 1565 LeuYsHisSerAsp-----ThrGlnAlaGlnLeuGlnIuYsThrGlnGln 1580
 Db 22426 GTAAACAAATATGATGTTTCATTAATGAAGATTCAGCCACCAAGATGCTTAATCAT 22485
 QY 1581 GluLeuGlnLeuValLysAsnLeuAlaIleAlaSerAspAsnYsProIleThrGln 1600
 Db 22486 GCAATTCACACACAAAGATTTGATTAACGCT----- 22518
 QY 1601 GluYsGluThrSerAlaAspCysValHisProLeu-----GluIuYsIleLeu 1617
 Db 22519 -----CATCAACTATCATGATTAATAATCAAAATGAT 22551
 QY 1618 LeuLeuThrGluGluLeuHisGlnIuYsThrAsnGluGlnGluYsLeuLeuHisGluYs 1637
 Db 22552 CAAGCTATGTAATAATATCAAAACAGACATTAATGAT-----TTACACGGTACT 22599
 QY 1638 AsnGluLeuGlnIleAlaGlnValGluLeuYsCysGluValGlnHisLeuMetLysSer 1657
 Db 22600 AATAAAGCTATCAACAGATTA-----AAAGAA 22626
 QY 1658 MetIleGluSerLysSerSerLeuGluSerLeuGlnHisGluYsHisAspThr---Glu 1676
 Db 22627 GCTTCAGAACCACTACAAACCTTAATAGCTTCACCAAGCGGCAAAAGATGATTTTA 22686
 QY 1677 GlnGlnLeuLeuAlaLeuYsGlnGlnIleMetGlnValValThrGln-----GluYsLys 1694
 Db 22687 AATCATATTTTCGTCACACCAACAGACAGCAAGTGAAGAAATAATGCAAGTCTTAA 22746
 QY 1695 GluLeuGlnGlnThrHisGluHisLeuThrAlaGluValAspHisLeuYsGluAsnIle 1714
 Db 22747 CAATTAATAATATCAATGAAGACACTTAGAGATTCATATTCGCTAATATATGAATTTTA 22806
 QY 1715 GluLeuGluYsAsnPheYsAsnGluAlaGlnGlnIuYsThrThrLysGluGlnCysLeu 1734

QY 2376 CysLeuGlnGlnHisGlyArgTyrSerAspSerAlaSerGlnLeuLeuLysPheCys 2395
 Db 24910 AAATTGGATGATTAAATCAGCAACAAAGACGTCAATATCTAATTAATATCTACTGCT 24969
 QY 2396 Gln-----ile 2397
 Db 24970 AAAACGACAAAGACAGTCCCAAGCAATGAAAGCAAAATCAATTAATCATGCAATG 25029
 QY 2398 GlnPheLeuAsnGlnLeuLeuPheLysAlaAsnIleGlnSer-----Val 2414
 Db 25030 AAAGCACTTAACAAATATATAAATGCAAGATTAAGTTTGATGATGATGATTCAT 25089
 QY 2415 GlnAspAspPheSerGlnValGlnValPhe-----LeuAsnGlnValGlySer--- 2430
 Db 25090 AAGCAAGATCAACCTGMAAAGGCGGTATCAACAAAGCTATTAATCATGATTCATCAATC 25149
 QY 2431 -----ThrLeuGlnGlnLeuGlnHisLys 2439
 Db 25150 ATTCAATAGCAACAACAAATCTGAAATGCATCCACAGATACCAATAGCATTAATCATGAA 25209
 QY 2440 LysGlyPheMetGlnTyrPheGlnGlnPheGlyAspLeuHisValAlaLysLysLeu 2459
 Db 25210 CTCGAACAGCTCA-----AATACCTTACATGCTGAT---CAGAACTT 25251
 QY 2460 SerGlnGlyMetGlnGlnLysAsnArgArgIleAlaSerThrIleGlnLeuLeuThrLys 2479
 Db 25252 GCTCATGCAACAAAGATGCCGCTAATGTATTAATGCTCAATTCATCTT----- 25302
 QY 2480 ArgLeuLysAlaValAlaGlnSerLysIleGlnArgGlnIleThrValTyrLeuAsnGln 2499
 Db 25303 -----AATGCTGCACAGCTGAGTAAATGATTAATACGAATCA 25341
 QY 2500 PheGlnAlaLysLeuGlnGlnLysLysGlnAsnLysGlnLeuMetArgArgMetGln 2519
 Db 25342 ---AATGCTCAACACGCGGAAGTTGCCAAAGAACTAGATTAATGCTCAAGCTCTGAT 25398
 QY 2520 HisHisGlyProSerAlaSerValMetGlnGlnLysAsnAlaArgLeuLeuGlyIleLeu 2539
 Db 25399 AAAGCTATGAAACACTCAACACAGTAGTCTCATATAAATAATATATGAAACGATAGT 25458
 QY 2540 LysThrValGlnAspGlnSerLysLysLeuGlnSerArgIleLysMetLeu----- 2556
 Db 25459 AAAATTTTAATGAAGATTCATAAATATCAACAAATACGATACGATTAATGCTGATGCC 25518
 QY 2557 GlnAsnGlnLeuAsnLeuValLysAspAlaMetHisLysGlyLysValAlaIle 2576
 Db 25519 GAACAACTACTTAATCAGACAAACAAATCCAACTTA---GAACTTAATAAAGCTGATATT 25575
 QY 2577 LeuGlnAspLysLeuLeuSerArgAsnAlaGlnValGlnLeuAsnAlaMetGlnValLys 2596
 Db 25576 GTTAAGATATATGCTAGCT---AAGCAAAATACTATTGCGCGCAAGAAAACHTATCA 25632
 QY 2597 LeuThrLysLysGlnAspAsnLeuGlnAlaAla---MetLysGlnIleGlnAsnLeuGln 2615
 Db 25633 TAGCAAAATCAAAATGAAATGAAATTAACATATGATTAATCTTAATATGACAA 25692
 QY 2616 LysMetValAlaLysGlnAlaValProTyr-----LysGlnGlnIleAspAsnLeu 2632
 Db 25693 AAGCAATCTTAATAAGATATGATTTCTCAGCGACATTAAGAACTAAGATTAAACAAT 25752
 QY 2633 -----LysThrLysVal 2636
 Db 25753 CTGCAACAAGCTAAATCTCTGATGAAAGCAATCAATCACTGAATAAACTCAAGTA 25812
 QY 2637 ValLysIleGlnMetGlnLysIleLysTyrSerLysAlaThr----- 2650
 Db 25813 GTGATTACAGATCTACTTGGCTTAATTACACTGAAGCTTCGAGAGATTAAGAAAAAGAAAA 25872
 QY 2651 ---AspGlnGlnIleAlaTyrLeuLysSerCysLeuGlnAspLysGlnGlnGly----- 2667
 Db 25873 GTAGACCAAACTGATACATGCTCAAGGATTAAT---GATTAATAATAATGCTCAAT 25929

QY 2668 -----LeuArgArgLeuLysGlnLeuArgAla-----GlnAla 2680
 Db 25930 GTAAGTTAGATCAAGTAGACAGCAAGCACTAGCAAAATTAATCAAGCATCAAGAAACCTC 25989
 QY 2681 AspAsnAspThrThrValCysValProLysAspArgGlnLysAlaSerThrPheProVal 2700
 Db 25990 GATGCTGATCAGCCAGATTGAAGAGCTTAAGTTCTACTTAATCAACAA----- 26037
 QY 2701 ThrCysGlyGlySerGlyIleValGlnSerThrAlaMetLeuValLeuGlnSerGln 2720
 Db 26038 -----ATTGATCAATTAACATCATCTTAATCATCAACAAACA 26076
 QY 2721 LysAlaAlaLeuGlnArgGlnLeuSerHisTyrLysLysLysTyrHisLysSerArg 2740
 Db 26077 ACTCG-----AAGAAAGTGTAAAGCAACAAACAACTAGAAAGAAATCGCTACT 26127
 QY 2741 ThrMetSerSerGlnAspArgLysThrLysAlaLysSerAlaHisSerSer 2760
 Db 26128 GTTAGTAACAATGCTCAGCATTAACAAAGTAATGGTAAATTAACAATTCATTAAT 26187
 QY 2761 HisThrGlySerSerHisArgGlySerProHisLys----- 2772
 Db 26188 CATGCTGATTCCTGTGAATAATAGATTAATTAATAGCAAGCCGACGACAAATCATC 26247
 QY 2773 -----ThrGlnThrTyrArgHisGlyProValThrProGlnArgSerGlnMetProSer 2790
 Db 26248 GCTTATGATGACACACTGAACATGACAA----- 26277
 QY 2791 LeuHisLeuGlySerProLysLysSerGlnSerSerThrLysArgValAlaSerProAsn 2810
 Db 26278 ---GATATACAAAACCTAACCCCAACCAAAATGAACAAACAAACGCTTA----- 26325
 QY 2811 ArgSerGlnIleTyrSerGlnLeuValMetSerProGlyLysThrGlyMetHis----- 2828
 Db 26326 CAACATTAATATATATGACAAACATGCTTAATAGTTGCGAAAGATTAATATCATCTAGA 26385
 QY 2829 -----LysHisIleLeuSerProSerLysValGlyLeuHisLysLysArgAla 2844
 Db 26386 CCAGAGCTTTAGAAATATATCAATCAATCACTAGAAATAAACAAGCTCAAAAGCTGCT 26445
 QY 2845 LeuSerProAsnArgSerGln-----MetProThrGlnHisValIleSerPro 2860
 Db 26446 TTAGAGATTAAGTAAAGCAATGCGATGATTAATTAAGATTGAACATATGCTCAACAG 26505
 QY 2861 GlyLysThrGlyLeuHisLysAsnLeuThrGlnSerThrLeuPheAspAsnLeuSerSer 2880
 Db 26506 GGC----- 26508
 QY 2881 ProCysLysGlnGlnLysValGlnLysAsnLeuAsnSerProLysGlyLysLeuPheAsp 2900
 Db 26509 -----ACAAACCTCAATGACATTAATGCGTAATGCTAAC 26544
 QY 2901 ValLysSerLysSerMetProTyrCysProSerGlnPhe-----Phe 2914
 Db 26545 GCAATCGTTAATTAAC---TATGCTCCAAACCAAGCAAGTAAATTAATTAACGCC 26598
 QY 2915 AspAsnSerLysLeuGlyAspPheSerGlnLeu-----AsnThrAlaGlnLysAsnAsp 2932
 Db 26599 GATTAACCTACGCAAGAAAGATTAATCTTACTCAACCTATCAACATGACGAGTACCTCAAC 26658
 QY 2933 LysSerGlnAlaGlnAsnTyrPyrGlnAla 2943
 Db 26659 AAAACTCAAGCTCAAGACTTGAATTTCAATGCA 26691
 RESULT 14
 US-09-166-350-27
 ; Sequence 27, Application US/09166350A
 ; Patent No. 6440663
 ; GENERAL INFORMATION:
 ; APPLICANT: Scanlan, Matthew
 ; APPLICANT: Chen, Yao
 ; APPLICANT: Stockert, Elisabeth
 ; APPLICANT: Old, Lloyd

QY	1355	sergluylleugluLeuLeuAlaHisValIysSerValGluGluAsnLeuIle	1374
Db	1613	CCAGAACTTGAAATATACCAATAAAGAACTTCACAGAAAGATGAGTATACTTACTAGT	1672
QY	1375	ThrlYsIleuAsnGlyLeuGluIlyGluIleleuGlyLysSerGluIserGluVal	1394
Db	1673	CTCAGCTCAAAGAGATACCAAGCTTTAAAGATTTGAGAGAAAG-----	1714
QY	1395	LeuIysSerMetLeuGluAsnLeuLysGluAspAsnLysLeuYsGluGlnAlaGlu	1414
Db	1715	ATAAATTTCTCTTACTAGAGAAAAGATGATTTTAAATTAATCTGAAAAATTCCTCATGA	1774
QY	1415	gluYrSerSerIysGluAsnGlnPheSerLeuGluGluValPheSerGlySerGlnIys	1434
Db	1775	GAATAGGATTAATTCCTATCAAGAAATGTGMAAGGACAGAA-----ACA	1816
QY	1435	LeuValAspGluIleGluValLeuYsAlaGlnLeuIlysalAlaGlnGluIArgLeuGlu	1454
Db	1817	TTGATTTCTTGAACTTGG-----AAGAAAGTAGAGCAAAACAAATCCAG	1858
QY	1455	IleIysAspArgAspTrpPheGluLeuValGlnThrAlaAsnThrAsnLeuValGluIly	1474
Db	1859	TACAAACAGT-----GAACTTAGACAAAGATTAATTCGAA-----TTAACAGGA	1900
QY	1475	LysLeuGluThrProLeuAlaAspHisGlnGluIAspSerIleAspArgArgSerGlu	1494
Db	1901	GGACTAGAGAGACTTTAAA-----GAAAAGATCAAAATGACCAAAACCTAGAA	1951
QY	1495	gluMet-----GluIleLysValIleuGlyIlyLysGluIArgAsnGlnTrpLeuLeu	1512
Db	1952	AAACTTTGCTTCAAAAGAAAGTTCTCTCGAAGACAAAGAAATTTGTCAGCTGAATG	2011
QY	1513	GluIArgLeuGlnGluIlyLysLeuGluIserAsnLysLeuGluIleLeuGlnIysGlu	1532
Db	2012	AAGCTCTTTATAGAGAAACAAATTAACCTAGTTCAGAAAAAAACAGTTGAGTGGAT	2071
QY	1533	MetGluThrSerValIleuLeuLysAspAsp-----Leu	1543
Db	2072	TTGGGAGTTTTTTGTCTCAAAAAGAGATGTTATCTTAAAGACATATTCCTCAATTA	2131
QY	1544	GlnGlnIysLeuGlnIserLeuLeuSerGluAsnIleIleLeuLysGluAsnIleAspThr	1563
Db	2132	GAAGAAGAACTTCAGTTATATGCTTGAGAG-----CAAGCTAATTTTAAATTA	2179
QY	1564	ThrLeuYsHisHisSerAspThrGlnAlaGlnLeuGlnIysThrGlnGlnIleuGln	1583
Db	2180	CTGCTTGAAAGAGAC-----CAAGCTCAGAGATTAATTTGTTAAACTCG	2224
QY	1584	LeuAlaLysAsnLeuAlaIleAlaIAserAspAsnProIleThrGlnIlySGLu	1603
Db	2225	TTGTATGCTTTCTTTAAAGAAATGGAGTCAACAAGCTTTCAGAAAGCACTGAGAAAGAT	2284
QY	1604	ThrSerAlaAspCysValHisProLeuGluIlyLysIleLeuLeuLeuThrGluGluLeu	1623
Db	2285	GTT-----GTTAAATCTCTACAGCAGCTGCGTAATCTTGGCAAAAATTAATGAGAA--	2338
QY	1624	HisGlnLysThrAsn-----GluGlnGlnIlyLysLeuLeuHisGlnIlySAsnGln	1639
Db	2339	-----AAATGCAACTGGCTTTCAAGCTGATGAAAAAGATTA-----GAG	2380
QY	1640	LeuGluGlnAlaGlnValIleuLeuYsCysGluValGluHisLeuMetLysSerMetIle	1659
Db	2381	TTTGAAAA-----GAGATTAAGCTCTTCAAGAA-----	2410
QY	1660	gluIserLysSerLeuGlnIserLeuGlnHisGlnIlySGLuYsHisAspThrGluGlnGlnLeu	1679
Db	2411	GAGAGCTAGTCAAGTGTGAAGACTTAAGTCTTTATTCAGAGCACTAGAGCAAGAGAA	2470
QY	1680	LeuAlaLeuYsGlnGlnMetGlnValAlaThrGlnIlyLys-----	1694
Db	2471	GTTCCTTTAAGAAAGGTAGAAAGAAATACAGTCAAGAAAAAGAGGCCCTGCAGTCTGAT	2530

[illegible]


```

Db 3335 -----GTCAACTTCAAAAGAAAAGAAACAGCTTCAGAAAACCATGCAGAAATTA 3385
Oy 2060 LysAsnLeuThrAlaLysIleSerSerLeuGluGluIleLeuGlnAsnAlaSerIle 2079
Db 3386 GAG-----CTGCTTAAAAAAGATGGCCCAAAACCATGCAGATG 3421
Oy 2080 LeuAsnGluAlaValSerGluArgGluAsnLeuArgHisSerLysGlnGluValSer 2099
Db 3422 ATGAATATGCAAAATAGCTGATTATGAA-----CGTTTGATGAAA 3460
Oy 2100 GluLeuGluGluLeuSerLeuThrLeuLysSerArgAspHisAlaPheAlaGlnSerLys 2119
Db 3461 GAACTTAATATCA-----AAGTTAACTAATAATAAAC----- 3490
Oy 2120 ArgGluLysAspGluAlaValAsnLysIleAlaSerLeuAlaGluIleLysIleLeu 2139
Db 3491 -----ACACAGATACAGATTTGGAGCAAGAAATATAAATTCAA 3529
Oy 2140 ThrLysGluMetAspGluPheArgAspSerLysGluSerLeuGlnGluSerSerHis 2159
Db 3530 -----AAACAGAAACAGAAACCCCTACAGAAAGAAATTAATCTCA 3568
Oy 2160 LeuSerGluGluLeuGluCysThrLysThrGluLeuGlnMetLeuLysGlnGlnLysGlu 2179
Db 3569 TTACAGCTTCAGTACAAACATATGAGAAAGAAAACCAACAAATACACAAATGCTGTG 3628
Oy 2180 AspIleAsnAsnLysLeuAlaGlu---LysValLysGluValAspGluLeuLeuGlnHis 2198
Db 3629 AAAACCAAAAAGAACTGGCAGATTCACAAACAGACAGAAACAGTATCATTAATCTTCA 3688
Oy 2199 LeuSerSerLeuLysGluGlnLeuAspGlnIleGlnMetGluLeuArgAsnGluLysLeu 2218
Db 3689 ---GCATCTTAAAGGTGAGCTGGAGCCAGACCAAGCAGCAGTAGTAGTCTATTAATA 3745
Oy 2219 ArgAsnThrGluLeu---CysGluLysMetAspIleMetGluLysGluIleSerValLeu 2237
Db 3746 CAGCTGGCTGAATATACATCAGAGAACCAAAATCCAGACAC----- 3790
Oy 2238 ArgLeuMetGlnAsnGluProGlnGlnGluLysAspValAlaGluArgMetAspIle 2257
Db 3791 -----CTGAACACCTCTGCCGGAACACACAC-----CGTACG 3823
Oy 2258 LeuGluSerArgAsnGlnGluIleGlnGluLeuMetGluLysIleSerAlaValLysSer 2277
Db 3824 CTAAAGTCATACACAGCAGAGGTGACAGACACTACAGAGAGAGTCCGCTGCCAAGCA 3883
Oy 2278 GluGlnHisThrLeuLeuSerSerLeuSerSerGluLeuGlnLysGluThrGluAlaHis 2297
Db 3884 GAACAGACTACTGTAACTCT-----GAATTCGAGAGCTAC 3919
Oy 2298 LysHisCysMetLeuAsnIleLysGluSerLeuSerSerThrLeuSerArgSerPheGly 2317
Db 3920 AAA-----GTCCGAGTTCAATATGTTCTAATAACCAACAGAAAATAATCTATGTCT 3970
Oy 2318 SerLeuGlnThrGluHisValLysLeuAsnThrGln---LeuGlnThrLeuLeuAsnLys 2336
Db 3971 CAGGCTGAACCTGAGGCGCTAAACAGAAAGGAGACATCTGGAAATCTGATTGACAG 4030
Oy 2337 PheLysValValLysArgThrAlaAlaValLysGluAspHisSerLeuIleLysAspThr 2356
Db 4031 CTAAATAATC-----AAATTACAGATACGCAAAATTAATCTACAGATTAAAT 4075
Oy 2357 GluLysAspLeuAlaAlaGluGlnLysArgHisAspGluLeuArgLeuGlnLeuGlnCys 2376
Db 4076 GTATCTGAACCTCAACATTCGACATCTGACATCTGATACACG----- 4117
Oy 2377 LeuGluGlnHisGlyArgLysTrpSerAspSerAlaSerGluLeu----- 2392
Db 4118 CTAAAGAGGCAACAAATATCTCTCAGAGAACTGTCTCAAGAGCGCAACTCCGGGAA 4177
Oy 2393 LysPheCysGluIleGluPheLeuAsnGluLeuLeuPheLysAlaAsnIleIleGln 2412

```

```

Db 4178 AAATGTGTTCAATACAG-----TCAGAGAACATGATGTAATCTGACATACACAG 4231
Oy 2413 SerValGlnAspAspPheSerGluValGlnValPheLeuAsn----- 2426
Db 4232 ACTGTGATCAGCTTAACATCCAGAACGAGGCTCTTCGAATATAGCTTCCGAGATCAAGTG 4291
Oy 2427 -----GlnValGlySerThrLeuGlnGluLeuGlnHis 2438
Db 4292 CGACATTTCCAGAGAAACACAGAAAGACATGGAGACATTCACACAGACGCTCTCAAG 4351
Oy 2439 LysLysGlyPheMetGlnTrpLeuGlu-----PheGlyAspLeu 2452
Db 4352 ATGGAACACAGCTCTTCACGCTTAAGATGAACGACACAGAACGCCAGCTTCTCT 4411
Oy 2449 ----- 2448
Db 4412 CAACAACTTTGAGAACCTTCGAGAAAGAGAAACACAGACCTCCCTCTACACATG 4471
Oy 2453 HisValAspAlaLysLeuSerGluGlyMetGlnGlnLysAsnArgArgIleAlaSer 2472
Db 4472 CACACTGTAAACCCGGGAAGAGGAGAGACATGACAGACATGATACGAGTCTGTCT 4531
Oy 2473 ThrIleGlnLeuLeuThrLysArgLeuLysAlaValGlnSerLysIleGlnArgGlu 2492
Db 4532 TCCCGCAGCACATACACACAGCTTTAGAGCAGTG----- 4567
Oy 2493 IleThrValLysLeuAsnGlnPheGluAlaLysLeuGln-----GluLysLys 2508
Db 4568 -----CTTAACCTCTCCGAACTAACTGAGCCCTCATTAATGCACTGAA 4615
Oy 2509 GluGlnAsnLysGluLeuMetArgArgMetGlnHisHisGlyProSerAlaSerValMet 2528
Db 4616 TTACCAAGAAAGAAATGTGTTAGAGCTCAGAGCTTCCACACAAAAGTGCAGAT----- 4669
Oy 2529 GluGluGluAsnAlaArgLeuLeuGlyIleLeuLysThrValGlnAspGluSerLysLys 2548
Db 4670 -----CACTTAACCGCTCTGCTCGGGAACAGAAACCAACCAATCAAT 4714
Oy 2549 LeuGlnSerArgIleLysMetLeuGluAsnGlnLeu---AsnLeuValLysAspAspAla 2567
Db 4715 CTATAGAGCAATTAATGCTTCAAAAGTGAAATAGAAAGATGGAAGATCAAGAG 4774
Oy 2568 MetHisLysGlyGlu-----LysValAlaIleLeuGlnAspLysLeu 2581
Db 4775 CGAGAGAACTGCAGCGTAACCTGGAATACTGMAAGACGTCTCTGAGATCTATTTC 4834
Oy 2582 LeuSerArgAsnAlaGluAlaGlu-----LeuAsnAlaMetGln-ValLys 2596
Db 4835 TTGAACCAAGTATGAAAGAGAGAGACTTCTCTGTATTAATAACGATCTTGACAGCTC 4894
Oy 2596 sLeuThrLysLysGlnAspAsnLeuGlnAlaAlaMetLysGluIleGluAsnLeuGlnLys 2616
Db 4895 AGCCCTGAAGAAAGGAAACCTCTCGGCTGCTCAAGGTGAGAGAAAGAAATGCTTCC 4954
Oy 2616 sMetValAlaLysGlyAlaValAlaProLysLysGluLysLeuAsp----- 2630
Db 4955 CGTCTTCTGATGGCATCTCATCTTCATAGTGTGCTGAGCTCGATAGCTTATGTA 5014
Oy 2631 -----AsnLeuLysThrLysValValLysIleGluMetGln 2642
Db 5015 AGGAATATTTTATTAACCAATATAGATCTATTCAAAAAGTGTGACGTATATTACA 5074
Oy 2642 uLysIleLysTrpSerLysAlaThrAspGlnGluIleAlaLysLeuLysSerCysLeu 2661
Db 5075 CAAT-----CTTTTGCAAAAAGTGTGA 5099

```

RESULT 15
 US-09-724-517-1
 : Sequence 1, Application US/09724517
 : Patent No. 6379941
 : GENERAL INFORMATION:
 : APPLICANT: Berand, Christophe
 : APPLICANT: Freedman, Richard

; TITLE OF INVENTION: No. 6379941el motor proteins and methods for
 ; FILE OF INVENTION: their use
 ; FILE REFERENCE: 1031
 ; CURRENT APPLICATION NUMBER: US/09/724,517
 ; CURRENT FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: US/09/641,807
 ; PRIOR FILING DATE: 2000-08-17
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 3837
 ; TYPE: DNA
 ; ORGANISM: Human
 ; FEATURE:
 ; NAME/KEY: misc.feature
 ; LOCATION: (1226)...(1337)
 ; OTHER INFORMATION: n = a, t, c, or g
 US-09-724-517-1

Alignment Scores:
 Pred. No.: 1,48e-58 Length: 3837
 Score: 841.50 Matches: 342
 Percent Similarity: 42.40% Conservative: 244
 Best Local Similarity: 24.75% Mismatches: 537
 Query Match: 5.70% Indels: 259
 DB: Gaps: 47

US-09-150-867-1 (1-2954) x US-09-724-517-1 (1-3837)

QY 7 ValLysValCysValArgValArgProLeuIleGlnArgGluGlnGlyAspGlnAlaAsn 26
 DB 19 GValAAAGTTCCTTAAGAAATAGACCTCTTGCACAAAGACCTCTCATTAATCATCA 78
 QY 27 LeuGlnTrpLysAlaGlnAsnThrIleSerGlnValAspGly---ThrLysSerPhe 45
 DB 79 GTTGTGCTGAGATTAATCCAAACAGCCAGCAATTAATCTTTGGAGAGATAGAGTTTC 138
 QY 46 AsnPhesprArgValPheAsnSerHisGluSerThrSerGlnIleTrpGlnIleAla 65
 DB 139 ACTTTGATTTGTTTGTTCGCAAAATTCACCAAGATGAGTTTATACACATGTATA 198
 QY 66 ValProIleLeuArgSerAlaLeuGlnGlyTrpAsnGlyThrIlePheAlaTrpGln 85
 DB 199 AAGCCCTAGCTGTCTACTCATCATGAGGCTATATGCAACTGTCTTGGCTATGGACAA 258
 QY 86 ThrSerSerGlyLysThrTrpThrMetMetGly-----ThrProAsn 99
 DB 259 ACTGAGTCTGGGAAGACATACCATTTGGAGGGGCCATATGCTTCAATGTGGAGGCG 318
 QY 100 SerLeuGlyIleIleProGlnAlaIleGlnValPheLysIleIleGlnIlePro 119
 DB 319 CAAAGGGTTCATCTCTCAGCATTCACAGAAATATTTCAAGACATCTGTGAACATCCT 378
 QY 120 AsnArgGluPheLeuArgValSerTrpMetGluIleTrpAsnGluThrValLysAsp 139
 DB 379 AGCATTTGACTTATGAAAGATCTTATATGAAAGTGAAGAGAACCTTAAGAGAT 438
 QY 140 LeuLeuCysAspArgArgLysLysProLeuGlnIleArgGlnAspPheAsnArgAsn 159
 DB 439 CTCTTGAAATTTGAGACATCCATGAGATCTTCACATCCGAGAGAAAGTGAAGAAAC 498
 QY 160 ValTrpValAlaAspLeuThrGluGlnLeuValMetValProGlnHisValIleGlnTrp 179
 DB 499 ACAGTATTTGTTGGGCGCAAGATGCGCATGTGAGAGTGCAGGTGAAGATGAGTCTT 558
 QY 180 IleLysLysGlnLysLysAsnArgHisTrpGlyGluThrLysMetAsnPhisSerSer 199
 DB 559 TTGGATGATGGGAGATGAGCAGACATACAGTACCACTCAATGAATGAGCAGCTCAGC 618
 QY 200 ArgSerHisThrIlePheArgMetIleValGlnSerArgSprArgAsnProThrAsn 219
 DB 619 AGATTCATGCAATTTTACATCAGCATTTGTCAAGTTCATATAAATATATGAGGAGCT 678

QY 220 SerGlnAsnCysAspGlyAla-----ValMetValSerHisLeuAsnLeu 234
 DB 679 GAA-----GATGATCATGTGATATCCCTCGGACATATGTGTCGCAAGTTCACATT 729
 QY 235 ValAspLeuAlaGlnSerGlnArgAlaSerGlnThrGlyAlaGluValAlaGluLys 254
 DB 730 GTGATTTGGCAGATCAGAAAGTAAAGTAAACCAAGGGGAATCTGCTGAAGGTTCCAA 789
 QY 255 GlnGlyCysAsnIleAsnArgSerLeuPheIleGlnGlyGlnValIleLysLysLeuSer 274
 DB 790 GAATTCATTAATTCATATGATGATTCCTGCTTGAAGAAATGTATTAACCGCTCTGGG 849
 QY 275 Asp---GlyGlnAlaGlyPheIleAsnTrpArgAspSerLysLeuThrArgIleLeu 293
 DB 850 GACCACGCGAAGAGATTCACATATTCATATAGGATGCTAAATTAACCGGCTCTCG 909
 QY 294 GlnAsnSerLeuGlyGlnAlaLysTrpValIleIleLysThrIleThrProValSer 313
 DB 910 AAAGATTCCTGGAGGAGCTGTAAGCTGATGATGATGATGATGATGATGATGATGATG 969
 QY 314 -----PheAspGlnThrLeuSerThrLeuGlnPheAlaSerThrAlaLysHisValArg 331
 DB 970 TCGAATTTGATGATGCTTAATTTCTCAATATGCCAAAGAGCAGGACGAACTTAGA 1029
 QY 332 AsnThrProHisValAsnGluValLeuAspAspGlnAlaLeuLysArgTrpArgLys 351
 DB 1030 AACAAACCCCATGCTAACTCAGCCCCGAGCTCAGACCGT---ATAGATGAATGAATTT 1086
 QY 352 GlnIleLeuAspLeuLysLysGlnLeuGlu----- 361
 DB 1087 GAGATTAATTTGCTTGCAGAGCTTTGCAAGCCAGCAGGCTGCTGACCCAACTACC 1146
 QY 362 -----AsnLeuGluSerSerSerGluThrLysAlaGlnIleAlaLysGlnGlnHis 379
 DB 1147 CAGATCAATGAGAGAGGATCCTGATACAAATAGAGATTATCTCTGTAGAGCAAGTA 1206
 QY 380 ThrGlnLeuAlaGluIleLysGlnLeuHisLysGlnArgGluAspArgIleTrpHis 399
 DB 1207 GCTCAGCTTCAAGAGATGATNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1266
 QY 400 LeuThrAsnIleValAlaIleSerSerGlnGln----- 412
 DB 1267 NNN 1326
 QY 413 -----GlnAspGlnArgValLysArgLysArgValThr----- 424
 DB 1327 NNN 1386
 QY 425 -----TrpAlaProGlyLysIleGlnAsnSerLeuHis 435
 DB 1387 TTTGATACATTTGGGCGCATTAATTAATTAATCCATCAAGA---CAAGATTCAGAGAG 1443
 QY 436 AlaSerGlyValSerAspPheAspMetLeuSerArgLeuProGlnLysAsnPhisLys 455
 DB 1444 GTCCACACAAAGTCCGCTATGTAATCTCTGATCGAATATTTCTGTGATTTCCAAACGA 1503
 QY 456 AlaLysPheSerAspMetProSerPheProGlnIleAspAspSerValCysThrGluPhe 475
 DB 1504 AGTCAGATGCTG---TTGGTCAATAGAGACAGATGAAGTCTCCATCCCAATTT 1560
 QY 476 SerAspPheAspArgAlaLeuSerMetMetAspSerHisGlnIleAspAlaGluTrpAsn 495
 DB 1561 TCTGATTAACAGTATGAT-----GAAGAATGAGAGGCCAAAGAAATCTGGAACCT 1611
 QY 496 LeuAlaSerLysValThrHisArgGluLysTrpSerLeuHisGlnSerMetIleAspPhe 515
 DB 1612 AGATGTAAGATGCTTCATGAGATTCAGAAACCAAGCTCTGTTTCTGTTGAATGG 1671
 QY 516 GlyGlnIleSerAspSerValGlnPheHisAsp----- 526
 DB 1672 AGTGAATCTCAGATGAAACACAAAGTCAATTTGGAGAAATGAATTAAGATTTGAT 1731
 QY 527 -----SerSerLysGlnAsnGlnLeuGlnTrpLysAspSerGlyAspMet 543

Db 1732 TGTCTCCAGAGAGTCAGAAATTTGCAAAAATTAAGAAATTCAGAACGATCACTT 1791
 QY 544 AlagluCysArgLysAlaSerPheGluLysGluIleThrSerLeuGlnGlnLeuGln 563
 Db 1792 ACTGAGGTAAA-----CAAAAATGAGAGAACTTCAATTAACATCAAG 1836
 QY 564 SerLysGluGlnLysLeuValGlnSer----- 575
 Db 1837 ATGAAGGAGATCTGATTAAGAAATTAATAAAGAGTAAATGATCAAGCTCTAAGC 1896
 QY 576 -----PheGluLeuLysIleAlaGluLeuGlu-----GluGlnLeuSerValLys 590
 Db 1897 AAGCAGTATCTTGAAGTAACAAGCTAGCATGATGCAAGACAGCAAAAGCTGAA 1956
 QY 591 -----AlaLysAsnLeuGlnMetValThrAsnSer-----ArgGluHisSer 604
 Db 1957 CTGATGTAACACAAACACACCTACAGAGCTGAAACAAAGATCTTCTGATGTTCA 2016
 QY 605 IleAsnAlaGluValGlnThrAspValGluLysGluValValArgLysGluMetSerVal 624
 Db 2017 ATGAAGGTAATTAACAAAGAGATTGCTTAAGATGATGCTGCAAAAGCTGAGAGTT 2076
 QY 625 LeuGlyAspSerGlyTyrAsnAlaSerAsnSerAspLeuGlnAspSerSerValAspLys 644
 Db 2077 CAGGCTTCGAGAGAACACACAGATAGTAAGAACTGGCATCCTGTCATCCAAAT 2136
 QY 645 LysArgLeuSerSerSerHisAspLysIleGluHisArgLys-----Met 660
 Db 2137 GAGAAACCTGCTAATGACCTGAGCAGAGTGTAGTACATGAAATATCAAAAGATACAG 2196
 QY 661 LeuGluGlnLysIleValAspLeuGlnLupheIleGluAsnLeuAsnLysLeuSerGlu 680
 Db 2197 CTCAAGAAAGAACTACGAGAAAGAAATGAAAAAGAACCAAGCATGATGCAATTAAG 2256
 QY 681 AsnAspLysGlnLysSerSerGluGlnAspPheMetGluSerIleGlnLeu----- 697
 Db 2257 CGGAGCACGCAAAAATCAAGAA-----ATCAATTAATAAAGACAGA 2298
 QY 698 CysGluAlaIleMetAlaGluLysAlaAsnAlaLeuGlnLupheAlaLeuMetArg--- 716
 Db 2299 CAGAGAAAGGCTTAAACCGAAGCTGAGACCTGTGATGATTAACATAAAGGAGA 2358
 QY 717 -----AspAsnPheAspAsnIleIleLeuGlnAsnGluThrLeuLysArgLys 732
 Db 2359 AAAGTTCCTTGAAGATATGACCATCTCCAGAAATTGATGACCAAAAGAAAGGTTA 2418
 QY 733 IleAlaAspLeuGluArgSerLeuLysGluAsnGlnGluThrAsnGluPheGlu----- 750
 Db 2419 GATGAGAGATGAGAAAGTTCTGAACCAACGCCAAGATTAAGAGACCTGGAACAGAC 2478
 QY 751 IleLeuGlnLysGluThrGlnLysGlnHisGlnLupheIleHisGluIleLeuLys 770
 Db 2479 TTAAGAAACCGAGGCCATAGTTCTTAAGAAAGAGCTCTGTTACAGAGAAAGAGTAC 2538
 QY 771 LeuLys---LysLeuValGluAsnAlaGluMetLysTyrAsnGlnAsn----- 784
 Db 2539 CTGAAAAATAGAAATAGAGTCTAGTACAGCCCTTAACACAGATAGTTGAAAAATATCA 2598
 QY 785 -----LeuGluGlnAspLeu----- 789
 Db 2599 ACTGCGCTGAACCTTACTGGAACAGAGTTCTGTAAGAAAGATGTCAGACCTCAGACCA 2658
 QY 790 -----GluThrLysThrLysLeuLeuLysGluGlnGluIleGlnLeuAlaGluLeuArg 807
 Db 2659 ACAGCTGGGAGAAACAAATTTCAACACAGTGAAGTC-----CTCAG 2706
 QY 808 LysArgAlaAspAsnLeuGlnLysValArgAsnPheAspLeuSerValSerMetLys 827
 Db 2707 AAAGAAAAAGATCAAGCTCCAAAGAGCAGACAGATGTGATGAAGAACTTAAGAAATGT 2766
 QY 828 AspSerGlnLysLeuCysGluGlu-----IlePheGlnLeuLysGlnSerLeuSerAsp 845

Db 2767 AGAGTGTATACACTGGAAGAAACATGTTCTTCCAACTTGAAGAAAGGATAGACCT 2826
 QY 846 AlagluAlaValThrArgAspAlaGlnLysGluCysSerPheLeuArgSerLupheAsn 865
 Db 2827 TTGGAAGCTGCATTTGAA-----TACAGGAATGAAGATATC 2862
 QY 866 GluLeuLysGluLysMetGluAspPheSerAsnTyrAsnGlnLysGluLysAlaIle 885
 Db 2863 CAGAAATCCGCAAGATGACTTACAGCATCA---TTCATTAACCTCTCTGCTGACAGCA 2919
 QY 886 SerLeuPheGluLys-----GlnLeuGluThrGluLysSerAsn 898
 Db 2920 AATGCTTGGAAAGACTAGCTTGCCTGAGTCTGTTGAGATTAAGACATTTCTTTACA 2979
 QY 899 TyrLysLysMetGluAlaAspLeuGlnLysGluGlnSerAlaPheAsnGluLeuAsn 918
 Db 2980 TAT-----TTCATTAAGTGTGTG 2997
 QY 919 TyrLeuAsnGluLeuLeuAlaGluLysValProArgAspLeuSerArgValGluLeu 938
 Db 2998 AATTGGCA-----GAGCT 3012
 QY 939 GluLysLysValSerGluPheSerLysGlnLeuGlu---LysAlaLeuGlnLupheAsn 957
 Db 3013 GAACGGAACACAGATTAATATGAAGAAATGAAGAAATGTAAGTCTGGAACGGCATAT 3072
 QY 958 AlaLeuGluAsnGluValThrCysLeuSerGluTyrLysPheLeuProAsnGluValGlu 977
 Db 3073 ATGTTCTGTAATTAAGATCTGACCTGACATCTAATA-----TTGCGAG 3117
 QY 978 CysLeuLysAsnGlnIleSerLysAlaSerGluGlnIleMetLeuLeuLysGlnGlu 997
 Db 3118 TGT-----GACCGAGAGCTACCCCTCCAGCAAAAG 3147
 QY 998 GluHisSerAlaSerIleLeuSerLysGlnLupheIleMetGlnGlnLupheSerGluGln 1017
 Db 3148 GAAACAGAA-----CAAAAGATGCGAGTCTATTAATCAATTCATTAACAAAGACAA 3195
 QY 1018 IleLeuGlnLeuThrAspLeuValThrHisThr---GlnSerLysValGlnGlnThrGlu 1036
 Db 3196 GATGAGAGAGGATTAATGAGAACTTCAAAACATTAATGAAGATTAATCCAGCAGTTGGA 3255
 QY 1037 GluGlnTyrLeuGluMetLysLysMetHisAspAsp-----Leu 1049
 Db 3256 AAAGATCTTATTTCTATTAAGAAACACCGCGATCATTAAGAAACTTAAGAACATG 3315
 QY 1050 PheGluLysTyrIleArgAsnLys-----SerGluAlaGluAspLeu 1063
 Db 3316 GTAGCGGAAGCAATTCGGCGGCACTAGCACCATCAGATTAACAAGAGCTGAGATGGA 3375
 QY 1064 LeuArgGluMetGluAsnLeuLysGluThrMetGluSerValGluValLysIleAlaAsp 1083
 Db 3376 GTCTGAAAGCCAGAGGA-----GAGGCGATGCTTTCGAAGAAATTAATAAGGCGATCC 3429
 QY 1084 ThrLysHisGluLeuGluGluThrIleArgAspLysGlu-----GlnLeu 1098
 Db 3430 AGACCTGAAGATGAATTAAGTGGAGAGAAAGAAAGAAAGAGTTCACAGACAGCAGC 3489
 QY 1099 LeuHisGluLysLysTyrPhePheGlnAlaMetGlnThrIlePheProIleThrProLeu 1118
 Db 3490 TTAAGAACACAGCCAAATCTCTAAAGCTCTGGGAGAGATATCCAGAAATTAACCTCAAT 3549
 QY 1119 SerAspSerLeu---ProPheSerLysLeuValGluLysAsnSerGluAspProIleGlu 1137
 Db 3550 CATAGCTCTTTAGCACCCCGCAGTGGCATATAGTTAGTAAGAAAT----- 3597
 QY 1138 IleAsnAspTyrHisAsnLeuIleAlaLeuAlaThrGluArgAsnAsnIleMetValCys 1157
 Db 3598 -----AAAACAGAAACAGATTAACATGATTACA 3627
 QY 1158 LeuGlnThrGluArgAsnSerLeuLysGluGlnValIleAspLeuAsnThrGlnLeuGln 1177
 Db 3628 AAATCTCACAGTCGACTCATCCCAAAATTCAGTTGTGGGAATGTGGAGACACTTCAT 3687

```
QY 1178 SerLeuGlnAlaInserIleGluYsSerAspLeuGlnLysProLysGlnAspLeuGln 1197
      ::      ::      ::      ::      ::      ::      ::      ::
Db 3688 GGTGTCACTGCTGTAATAACTGTGTGCGAAAAAGATTCGTAATTC----- 3735
QY 1198 GluGlyGluValLysLeuLeuGlnMetGluLeuLysGlyHisLeuThrAspSer 1217
      ::      ::      ::      ::      ::      ::      ::      ::
Db 3736 -----GCCTTGGAACATATCATTCGACGCTCCAGCTCT----- 3768
QY 1218 GlnLeuSerIleGluLysLeuGlnLeuGlnAsnLeuGlnValThrGluLysLeuGlnThr 1237
      ::      ::      ::      ::      ::      ::      ::      ::
Db 3769 GGACTTGGCATGTGATCAATGGCTGCTGATTCATCGAAGATCTAGGAAACCAAGGAC 3828
QY 1238 LeuGln 1239
      ::      ::      ::      ::      ::      ::      ::      ::
Db 3829 TTAATA 3834
```

Search completed: May 23, 2003, 15:16:20
Job time : 824 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 23, 2003, 15:03:25 ; Search time 709 Seconds
(without alignments)
5501.616 Million cell updates/sec

Title: US-09-150-867-1

Perfect score: 14769
Sequence: 1 MSECDAKVCVRRPRLQRE.....QAENWYEAKKETAPCKTS 2954

Scoring table:

BLOSUM62			
Xgapop 10.0 ,	Xgapext 0.5		
Ygapop 10.0 ,	Ygapext 0.5		
Fgapop 6.0 ,	Fgapext 7.0		
Delop 6.0 ,	Delext 7.0		

Searched: 828747 seqs, 660231138 residues

Total number of hits satisfying chosen parameters: 1657494

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL-frame_plus_p2n.model -DEV-x1p
-O/cgnt2_1/USPTO.spool/US09150867/runat_21052003_165914_21023/app_query.fasta.1.3143
-DB-published.Applications_NA-QFMT-fastap -SUFFIX-rnph -MINMATCH-0.1
-LOOPCL-0 -LOOPEXT-0 -UNITS-bits -START-1 -END-1 -MATRIX-blosum62
-TRANS-human40.cdi -LIST-45 -DOCALIGN-200 -THR SCORE-pct -THR MAX-100
-THR MIN-0 -ALIGN-15 -MODE-LOCAL -OUTFMT-pio -NORM-ext -HAPSTZE-500 -MINLEN-0
-MAXLEN-2000000000 -USER-US09150867_ECGN_1_1.518 &runat_21052003_165914_21023
-NCPU=6 -ICPU=3 -NO_XLPTX -NO_MAP -LARGEQUERY -NEG_SCORES-0 -WAIT
-DISPLOCK-100 -LONGLOG -DEV-TIMEOUT-120 -WARN-TIMEOUT-30 -THREADS-1 -XGAPOP-10
-XGAPEXT-0.5 -FGAPOP-6 -FGAPEXT-7 -YGAPOP-10 -YGAPEXT-0.5 -DELOP-6 -DELEXT-7

Database : Published_Applications_NA:*

1:	/cgnt2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2:	/cgnt2_6/ptodata/2/pubpna/PTCT_NEW_PUB.seq:*
3:	/cgnt2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4:	/cgnt2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5:	/cgnt2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6:	/cgnt2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
7:	/cgnt2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8:	/cgnt2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9:	/cgnt2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10:	/cgnt2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11:	/cgnt2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12:	/cgnt2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13:	/cgnt2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14:	/cgnt2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	3615.5	24.5	7992 9	US-09-893-519A-140 Sequence 140, App
2	3613.5	24.5	8493 9	US-10-071-766-51 Sequence 51, Appl
3	1153	7.8	10300 10	US-09-960-253-145 Sequence 145, App
4	1152	7.8	10096 10	US-09-960-253-163 Sequence 163, App

5	1139	7.7	10190 10	US-09-864-864-292 Sequence 292, App
6	1139	7.7	10211 10	US-09-954-456-1153 Sequence 1153, App
7	1139	7.7	10211 10	US-09-967-768A-186 Sequence 186, App
8	1024.5	6.9	12313 9	US-10-171-311-7 Sequence 7, Appl
9	1024.5	6.9	12438 9	US-10-171-311-3 Sequence 3, Appl
10	1022.5	6.9	12337 9	US-10-171-311-5 Sequence 5, Appl
11	1022.5	6.9	12462 9	US-10-171-311-1 Sequence 1, Appl
12	998	6.8	11677 9	US-10-083-830-134 Sequence 134, App
13	900.5	6.1	6386 9	US-10-098-841-40 Sequence 40, Appl
14	891.5	6.0	7497 12	US-09-960-253-175 Sequence 175, App
15	891.5	6.0	7792 12	US-10-044-090-359 Sequence 359, App
16	888.5	6.0	3366 9	US-09-938-842A-2651 Sequence 2651, App
17	848	5.7	6773 10	US-09-864-864-336 Sequence 336, App
18	836	5.7	14800 9	US-09-954-456-1601 Sequence 1601, App
19	829	5.6	7065 9	US-09-991-466-115 Sequence 115, App
20	829	5.6	7065 10	US-09-874-923-115 Sequence 115, App
21	797	5.4	15231 10	US-09-917-800A-1505 Sequence 1505, App
22	795	5.4	10625 9	US-10-023-219-3 Sequence 3, Appl
23	795	5.4	10625 10	US-09-727-384-5 Sequence 5, Appl
24	786.5	5.3	6586 10	US-09-954-456-1115 Sequence 1115, App
25	786.5	5.3	6586 10	US-09-954-456-1790 Sequence 1790, App
26	776	5.3	8948 10	US-09-735-105-119 Sequence 119, App
27	776	5.3	8948 10	US-09-850-116A-119 Sequence 119, App
28	776	5.3	8948 10	US-09-897-778-119 Sequence 119, App
29	776	5.3	9588 10	US-09-954-456-1848 Sequence 1848, App
30	762	5.2	7596 10	US-09-954-456-2215 Sequence 2215, App
31	754	5.1	13957 10	US-09-782-378A-22 Sequence 22, Appl
32	754	5.1	13957 10	US-09-880-107-2284 Sequence 2284, App
33	721.5	4.9	6861 9	US-10-171-311-161 Sequence 161, App
34	721	4.9	6900 9	US-10-171-311-163 Sequence 163, App
35	721	4.9	8930 10	US-09-954-456-1805 Sequence 1805, App
36	720.5	4.9	9274 10	US-09-885-535-3 Sequence 3, Appl
37	719.5	4.9	5937 9	US-09-927-597-3 Sequence 3, Appl
38	718	4.9	4780 10	US-09-962-436-287 Sequence 287, App
39	718	4.9	18846 10	US-09-815-242-8898 Sequence 8898, App
40	708.5	4.8	5835 9	US-09-927-597-1 Sequence 1, Appl
41	702	4.8	5848 12	US-10-044-090-299 Sequence 299, App
42	699.5	4.7	4858 10	US-09-954-456-733 Sequence 733, App
43	695.5	4.7	10302 10	US-09-782-378A-23 Sequence 23, Appl
44	691.5	4.7	9477 10	US-09-815-242-8513 Sequence 813, App
45	687.5	4.7	31096 7	US-08-781-968A-59 Sequence 59, Appl

ALIGNMENTS

RESULT 1
US-09-893-519A-140
; Sequence 140, Application US/09893519A
; Publication No. US20030027243A1
GENERAL INFORMATION:
; APPLICANT: ANADYS PHARMACEUTICALS, INC.
; APPLICANT: THOMPSON, Craig
; APPLICANT: MOORE, Jeffrey
; APPLICANT: BUDDMAN, Ed T.
; APPLICANT: BRADLEY, John
; APPLICANT: DESILVA, Thamara
; APPLICANT: HARRIS, Sandra
; APPLICANT: KOMARNITSKY, Svetlana
; APPLICANT: MENDILLO, Marc
; APPLICANT: MOORE, Daniel
; APPLICANT: MCCOY, Melissa
; APPLICANT: SANDERSON, Karen
; APPLICANT: HAO, Tarig
; APPLICANT: ZHU, Shuhao
; APPLICANT: LONG, Fan
; APPLICANT: DAVIDOV, Eugene
; TITLE OF INVENTION: ANTIFUNGAL COMPOUNDS AND METHODS OF USE
; FILE REFERENCE: 0342/16548-US2
; CURRENT APPLICATION NUMBER: US/09/893,519A
; PRIOR FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 60/215,164
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: US 60/224,457

Qy 5318 ValArgLysGluMetSerValLeuGlyAspSerGlyTyrAsnAlaSerAsnSerAspLeu 637
1506 ----- 1506
Qy 638 GluAspSerSerValAspGlyLysArgLeuSerSerHisAspGluCysIleGluHis 657
1506 ----- 1506
Qy 658 ArgLysMetLeuGluGluHisIleValAspLeuGluGluPheIleGluAsnLeuAsnLys 677
1506 ----- 1506
Qy 678 LysSerGluAsnAspLysGluLysSerSerGluGluAspPheMetGluSerIleGluLeu 697
1506 ----- 1506
Qy 698 CysGluLalaIleMetAlaGluLysAlaAsnAlaLeuGluGluLeuAlaLeuMetArgAsp 717
1507 ----- 1515
Qy 718 AsnPheAspAsnIleIleLeuGluAsnGluThrLeuLysArgGluIleAlaAspLeuGlu 737
1516 GACTATGATATCTGATTAAGTACTATGACATGACACACTACAGACAGAAAAAGAAATGGA 1575
Qy 738 ArgSerLeuLysGluAsnGluGluThrAsnGluPheGluIleLeuGluLysGluThrGln 757
1576 TTGAATTAAGAAAAAGAAAGATTTGGATGATTTGGCGCTGAGAAAGAAACTAA 1635
Qy 758 LysGluHisGluAlaGluLeuIleHisGluIleGlySerLeuLysLysLeuValGluAsn 777
1636 AAAATCAAGAGATCAACCTATTCAGAAATTCGAAATTTTGTATAGCAT 1695
Qy 778 AlaGluMetTyrAsnGluAsnLeuGluGluAspLeuGluThrLysThrLysLeuLys 797
1596 CGAGAAAGATATTAATCAAGATCTTGAGAAATGACACTGATTAAGAAAGTACCTGTTAGA 1755
Qy 798 GluGluGluIleGluLeuAlaGluLeuArgLysArgAlaAspAsnLeuGluLysVal 817
1756 GAAAAAGAGACACAGATTAAGAACCTACAGAAATACATGACCTCCAAAGACTGGAAT 1815
Qy 818 ArgAsnAspLeuSerValSerMetGlyAspSerGluLysLeuGluIlePhe 837
1816 ATAAAAATGGACTTGTCTACTACTGTTG-----GAACCATTTGAGACCCAAA 1863
Qy 838 GluLeuLysGluSerLeuSerAspAlaGluAlaValThrArgAspAlaGluLysGluCys 857
1864 CAATGAGACGACTCTGTTGATGCTGAACCTGATGCCCTTGATGCCAAGAGAAATCA 1923
Qy 858 SerPheLeuArgSerGluAsnLeuGluLeuLysGluLysMetGluAspThrSerAsnTrp 877
1924 GCCTTCTTGAAGTGAATAATCTGAGTGAAGAGAAAAAGAAAGAACTGCAACTCA 1983
Qy 878 TyrAsnGluLysGluLysAlaIleSerLeuPheGluLysGluLeuGluThrGluLysSer 897
1984 TACAAAGCAAAATGGAAAAATGATATTCAGTTATATCAAAAGCCAAATTTGGAGGCAAAA----- 2037
Qy 898 AsnTyrLysLysMetGluAlaAspLeuGluLysGluLeuGluGluSerAlaPheAspGluIle 917
2038 -----AAGAAATGCAAGTTGATCTGAGAAATAATTAATATGCTTTAAAGAGATA 2091
Qy 918 AsnTyrLysAsnGluLeuLeuAlaGlyLysValProArgAspLeuSerArgValGlu 937
2092 ACAAAACTCACTCCCTTATATAGATGGCAAGTCCAAAGATTTGGCTGTGAATTTGGAA 2151
Qy 938 LeuGluLysLysValSerGluPheSerLysGluLeuGluLysAlaLeuGluGluLysAsn 957
2152 TTGGAGAGAAAGATTAAGTCTTCAAGAAAGAACTTAATAAGAAATTTGAAGAAAGAA 2211
Qy 958 AlaLeuGluAsnGluValThrCysLeuSerGluTyrLysPheLeuProAsnGluValGlu 977
2212 GCTTTGGGAGAAAGCATTTTGTCTTCAAGAAATGAAATCTTACCTTGTGAAGTGA 2271

Qy 978 CysLeuLysAsnGluIleSerLysAlaSerGluGluIleMetLeuLeuLysGluGlu 997
2272 AGCCTGAGAGAAAGATATCAAGACAAATCTGAAGGCTCCATATTAATATCACTGCAAAAA 2331
Qy 998 GluHisSerAlaSerIleIleSerLysGluGluIleIleMetGluGluGluSerGluGlu 1017
2332 GATAAATGTTTCTGAGATGATTCATAGAGAGAGATGAGATGCAAGCTTACTTGAGAA 2391
Qy 1018 IleLeuGluLeuThrAspGluValThrHisThrGlnSerLysValGluGluThrGluGlu 1037
2392 ATTGGAAGAAACAAAGACGACCTACCAACTACAGGTGCAATTAATTAAGCACTGATCA 2451
Qy 1038 GluTyrLeuGluMetLysLysMetHisAspAspLeuPheGluLysTyr----- 1053
2452 GAATTCAAAATTTCAAAACCTTCATATGACCTTGAGCAAAAGATTAATGATGCTCTT 2511
Qy 1054 -----IleArgAsnLysSerGluAlaGluAspLeuArgGluMetGluAsnLeu 1070
2512 GAGGAGAAATGAGAGAAATGAATCAGAGAAATGATTAATCTCTCAAAAGAACCCAAATTT 2571
Qy 1071 LysGluThrMetGluSerValGluValLysIleAlaAspThrLysHisGluLeuGluGlu 1090
2572 GATTCGAGTTGGGTGCTTTGAAGCCGAGCTTCTTCAAGACCCAAAGAACTTCAGAG 2631
Qy 1091 ThrIleArgAspLysGluGluLeuLeuHisGluLysLysTyrPhePheGluAlaMetGln 1110
2632 AAACACGTCGAGCTTCAAGAAAGACTAATGATGAGAGAACGTCGAAAGCAATTAAGAA 2691
Qy 1111 ThrIlePheProIleThrProLeuSerAspSerLeuProSerLysLysValGluGlu 1130
2692 AAT-----AGAGATTCCTCCCTCGCAAACTGAGAAAGGAGAAACACTGATTAAGT 2745
Qy 1131 AsnSerGluAspProIleGluIleAsnAspTyrHisAsnIleAlaLeuAlaThrGlu 1150
2746 AAACGCGACCAACTTTAGAA-----GAGTAAAACTTTAACTCAAGAAAGAAAGATAT 2799
Qy 1151 ArgAsnAsnIleMetValCysLeuGluThrGluArgAsnSerLeuLysGluGluVal 1169
2800 CTAAAGAACATCCAAAGAAAGCTTGCAATTTGAGAGGAGCAACAAAGATGATTTGAC 2859
Qy 1170 -----IleAspLeuAsnThrGluLeuGln-----SerLeuGluAlaGln 1182
2860 GATACGTTAATCATGATATATGATATCTCAAGAACAAATTCACAAATTCCTGAGACTCTG 2919
Qy 1183 SerIleGluLysSerAspLeuGluLysProLysGluAspLeuGluGluGluValLys 1202
2920 AAACAACTCAAGAAACATTAATACACTTAAATCGAAAAATTTCTGAG--GAGGTTTCC 2976
Qy 1203 LeuLeuLeuGluMetGluLeuLeuLysGluLysLeuThrAsp-----SerGln 1218
2977 AGGAATTTGCATATGAGAGAAATATACAGAGAAACTAAAGATGAATTCACGCAAAAGATG 3036
Qy 1219 LeuSerIleGluLysLeuGluLeuGluAsnLeuGluValThrGluLysLeuGluThrLeu 1238
3037 GTTGCAATGATTAATAAAACAG-----GATTTGGA--GCTAAATAATACCCAAACACTA 3087
Qy 1239 GluGluGluMetLysAsnIleThrIle----- 1247
3088 ACTGCAGATGTTAAAGGATATGAGATTAATGAGCAACAAAGAAAGATTTTCTTTAATA 3147
Qy 1248 --GluArgAsnGluLeuGluThrAsnPheGluAspLeuLysAlaGluHisAspSerLeu 1266
3148 CAGAGAGAAATGAACTCCAAACAAATGTTAGAGACTGTTATAGCAGAAAGAAAGCAATTTG 3207
Qy 1267 LysGluAspLeuSerGluAsnIleGluGluSerIleGluThrGluAspGluLeuArgAla 1286
3208 AAGACTGCTCAAGAAATATTTGAATGACCAATTTGAAACCGAGAAAGATTAAGACTT 3267
Qy 1287 AlaGluGluLeuArgGluGluGluLysGluLeuValAspSerPheArgGluGluLeu 1306
3268 CTGCGGATGAACTTAAGAAAGCAACAGATGATCTGCACAAAGAAAGCAATGCCATA 3327
Qy 1307 AspCysSerValGlyLysSerSerProAsnHisAspAlaValAlaAsnGluGluVal 1326

Db	3328	AAG----	AAAGAAAGAGAGCTTTCTAGAGACCTGTGACAGACTGGCAGAGTGCAGAAAGAAAA	3384	
Oy	1327	serLeuGluIyValAsnSerLeuGlnSerGluMet-----	LeuArgGly	1341	
Db	3385	CTAAAGAGAAAGAGAGCCAGCAACACTCCCAAMAAAACAGCAACACTCTTAAATGACAAAGA	3444		
Oy	1342	GluArgAspGluLeuGlnThrSerCysLysAlaLeuValSerGluLeuGlnLeuArg	1361		
Db	3445	GAGATGAGTGAAGTGCAGAAAAAG-----	ATTAAATGAATAAGCAATTTTAAAG	3492	
Oy	1362	AlaHisValLysSer-----	ValGluGlyGlnAsnLeuGluIle	1374	
Db	3493	AATGAATTAAGAAGCAAAAGATTGACATGTGACATATGGAAMAGAGAGAGCTTGAGTGG	3552		
Oy	1375	ThrLysLeuAsnGlnLeuGlnIuysGluIleLeuGlnLysSerGluGlnSerGluVal	1394		
Db	3553	GCTCAGAAACTTAATGAAAAATTATGAGGAGGAAAGTAAATCTTAACCAAGAAAGAAAAAGTT	3612		
Oy	1395	LeuLysSerMetLeuGlnAsnLeuLysGluLysAspAsnAsnLysLysGluGlnAlaGlu	1414		
Db	3613	CTAAAGAGAAATTACAGAAAGTCATTTGAAACAGAGAGAGACCACTTAAAGAGATAT-----	3666		
Oy	1415	GluTyrSerSerLysGlnAsnGlnPheSerLeuGluGluValPheSerGlySerGlnLys	1434		
Db	3667	-----	ATMAAGAAATTTGAAGCTACAGGCTTACAA	3696	
Oy	1435	LeuValAspGluIleGluValLeuLysAlaGlnLeuLysAlaIleGluGluArgLeuGln	1454		
Db	3697	ACCAAGAGAAAGAACTTAAAAATTTGCTCATATTACACTTAAAGAAACCAAGCAAACTATGAT	3756		
Oy	1455	IleLysAspArgAspTyrPheGlu--	LeuValGlnThrAlaAsnThrAsnLeuValGlu	1473	
Db	3757	GAACATAAGAAAGAACCTTATCTGAGAGACAGCTCAAAATTAATTAATACAGAGACTTAAAGAA	3816		
Oy	1474	GlyLysLeuGluThrProLeuGlnAlaAsp-----	HisGluGluAsp-----	1487	
Db	3817	---AAATCCCATACCAAAATTTCAGAAAGAGATCCCAAGCTTCATGAGAGAACAGATTA	3873		
Oy	1488	-----	SerIleAspArgArgSerGluGluMet--	GluIleLysValLeu	1501
Db	3874	CTGCCTAATGTGAAAAAGAGCTAGTGAGACTCAGCAAGCAACATTAATGAAGTGGAGTTTATA	3933		
Oy	1502	GlyGluLys---	LeuGluArgAsnGlnTyrLeuLeuGluArgLeuGlnGluIleLysLeu	1520	
Db	3934	ACAGAACAGCTCCACACCAACAGGACTCAACACACTGGCAAGAAATGGAAGAAGGCTC	3993		
Oy	1521	GluLysSerAsnLysLeuGluIleLeuGlnLysGluMetGlyThrSerValLeuLeuLys	1540		
Db	3994	AGGTTGAATGAAAAAATTTCAGAAAGAGTCAGAGAGAGATAAATCTCTAACCAAGGAAGA	4053		
Oy	1541	AspAspLeuGlnGlnLysLeuGlnSerLeuLeuSerGluAsnIleIleLysGlnGlnAsn	1560		
Db	4054	GACAACTTTAAACGATTAAGAAAGGCCCTTGAAGATTAACTATGACCAGCTGGAAGAACAT	4113		
Oy	1561	IleAspThrThrLeuLysHisIleSerAspThrGlnAlaGlnLeuGlnLysThrGlnGln	1580		
Db	4114	ATTAGAGAAACTTGG-----	GCTAAATATCCAGAGAGCTTCAAGC	4152	
Oy	1581	GluLeuGlnLeuAlaLysAsnLeuAlaIleAlaIleAspAsnAsnCysProIleThrGln	1600		
Db	4153	AAACAGAGAACAGACTCTTAATTAATGAAAGAAAGAAAGACAAATGAATCTGAGT	4212		
Oy	1601	GluLysGlu-----	ThrSerAlaAspCysValHisProLeuGlnLysLysLeu	1617	
Db	4213	GAGATGAGCAATTCAAACCCAAAGATTCAGCACTACTAAGATTAAGATTAATAAATGCTC	4272		
Oy	1618	LeuLeuThrGluGluIleHisGlnLysThrAsnGluGlnLysLeuLeuHisGluLys	1637		
Db	4273	GGAATTGTCCAAAAGACTTCAGAAAGATGATGAATTAAGAAATCTTACTAAGAGAGAAA	4332		
Oy	1638	AsnGluLeuGluGlnAlaGlnValGluLeuLysCysGluValGluHisIleuMetLysSer	1657		

[illegible]

QY 3081 GUAUAUAAnLysValIleAlaLeuThrGluLysMetSerSerLeuGluGluIle 2000
 Db 5407 AACTATCAAAATATGCAAAAAGATTAGAAAATTCAAATGCTAAATTTCAGAAAGATT 5466
 QY 2001 AsnGluAsnValThrThrLeuLysGluLysGluLysGluThrPheTyrLeuGln 2020
 Db 5467 CAAGAA-----CTTAAGGCAAAATGACATCAATCTATTACTTAAAAAGAT 5514
 QY 2021 ArgProSerLysGlnGlnSerSerGlnMetGluLeuArgGluSerLeuLysThr 2040
 Db 5515 GTCAATGAGACAGCAAAAAGATGCTGAAATGAGCAATCAAAAGAAATAAAC 5574
 QY 2041 LysAspLeuGlnGluGluAlaGluLysGluLysSerGluAlaThrAsnGluLys 2060
 Db 5575 CAAGCTTAAGCTGAGAAATTA-----GAAATAGAG 5607
 QY 2061 AsnLeuThrAlaLysIleSerSerLeuGluGluLysGlnIleAsnAlaSerLeu 2080
 Db 5608 AATTTA-----AATTGGCTCAAGAACTTCATGAATAAACCCTGAAGAAATG 5652
 QY 2081 AsnGluAlaValSerGluArgGluAsnLeuArgHisSerLysGlnGlnLeuValSer 2100
 Db 5653 AAATCTGTAATGAAGAAGATATCTAAGAAGAGTAGAGACACTAACTGAG 5712
 QY 2101 LeuGluGlnLeuSerLeuThrLeu-----LysSerArgAspHisAlaPheAlaGln 2117
 Db 5713 AGAACCAACCTCAAGAAAGCTCCAGAAACCAAGCTAGATCTGGAATACACAG 5772
 QY 2118 -----SerLysArgLys-----AspGluAlaVal 2136
 Db 5773 GAACATAAACTGCTGATCTATCAAAAGAACAAAGAACTGTGATGAATACTTGA 5832
 QY 2127 AsnLysIleAlaSerLeuAlaGluGluLysIleLeuThrLysGluMetAspGluPhe 2146
 Db 5833 GAAAAAATTCAGAAAGACAAATTCAAATTCAGACATCAAAAGGATTTGATGATAA 5889
 QY 2147 ArgAspSerLysGlnSerLeuGlnGlnIleSerHisLeuSerGluLysCysThr 2166
 Db 5890 -----TCAAAAGATGAATTCACAAAAAGATCCAAAGACTTCAG----- 5928
 QY 2167 TyrLysThrGluLeuGlnMetLeuLysGlnGlnLysGluAspIleAsnLysLeuAla 2186
 Db 5929 -----AAAAAGAACTTCAACTCTTAGA-----GTGAAAGAAATGTCATATGATG 5979
 QY 2187 GluLysValLysGluValAspGluLeu----- 2195
 Db 5980 AAAAAAATTAATGAATGAACAGTGAAGAAAGCAATTTGAGCCAACTATCTATGCAAG 6039
 QY 2196 -----LeuGlnHisLeuSerSerLeuLysGlnGlnLeuAspGlnIleGln 2210
 Db 6040 TGTGAGATGATTAAGTCCAGCTTACAGAACTTCATGAAAGCTTTGAAGAAATTAAGA 6099
 QY 2211 MetGluLeuArg--AsnGluLysLeuArgAsnTyrGluLysCysGluLysMetAspIle 2239
 Db 6100 ATTGTACTTAAGAAAGAGATGAGCTAGAGGATTAAGAAATCTCTCAAAATGGA 6156
 QY 2230 MetGluLysGluLysSerValLeuArgLeuMet-----Gln 2241
 Db 6157 AGGACCAATTCATAGCAACCTTAAGGAAATGTAGCTAGAGACCGACAGAACCCACCA 6216
 QY 2242 AsnGluProGlnGlnIle-----GluAspAspValAlaGluArgMetAspIleLeuGln 2259
 Db 6217 GTAAACCTGAAAAAGGTTACTAGTGAATGAGACACAGCACTTATGGAAGGCTTAGA 6276
 QY 2260 SerArgAsnGlnGluIleGlnGluLeuMetGluLysIleSerAlaValTyrSerGluGln 2279
 Db 6277 GAAAGAGCTCTAGAAATTAAGAGCTTTGGAAGAGATCTACAGAGATG--GATGATCAT 6333
 QY 2280 HisThrLeuLeuSerSerLeuSerGlnLeuGlnLysGluThrGluAlaHisLysHis 2299
 Db 6334 TATGAGCTTGAATTAAGATGCTCTTGAAGAGGAAATTAATTCACAGAA-- 6390

QY 2300 CysMetLeuAsnIleLysGluSerLeuSerThrLeuSerArgSerPheGlySerLeu 2319
 Db 6390 ----- 6390
 QY 2320 GlnThrGlnHisValLysLeuAsnThrGlnLeuGlnThrLeuAsnLysPheLysVal 2339
 Db 6391 -----ATCATGAAACAACTGAGAT 6411
 QY 2340 Val-----TyrArgThrAlaAlaValLysGluAspHisSerLeuIleLysAspTyrGlu 2357
 Db 6412 GTGTTAACTTATGTTCAAAAAATAAAGAACCAACATGTAATGCATTAATTAATTTGAA 6471
 QY 2358 LysAspLeuAlaAlaGluGlnLysArgHisAspGluLeuArgLeuGlnLeuGlnCysLeu 2377
 Db 6472 ATGATTTTATGATGAAGGTGGAAGCAAGCAATTCATTAATTAATTAATTAATTAAT 6531
 QY 2378 GlnGlnHisGlyArgLysThrSerSerAspSerLeuGluLysPheCysGluIle 2397
 Db 6532 CAACAAGAT-----TGATGTATACCATTCACAGCAATTAAGGATCTCAAAATG 6579
 QY 2398 GluPheLeuAsnGluLeuLeuPheLysLysAlaSerIleGlnSerValGlnAspAsp 2417
 Db 6580 AACCAAAATATGATCTA-----CATATTGAGAAATTCACAAAGAT 6621
 QY 2418 PheSerGluValGlnValPheLeuAsnGlnValGlySerThrLeuGlnGlnLeuGlu 2437
 Db 6622 TTCTCAGAAAGTAG-----TTCCCTAGCATTAAGACTGAATTCACAAAGCTACTAGT 6675
 QY 2438 HisLysLysGlyPheMetGlnThrLeuGlnGluPheGlyAspLeuHisValAspAlaLys 2457
 Db 6676 AATGGAAGAAATGACACACAGCTTTTGGAAAGCTGTAATTAATCTGTTGATATAGA 6735
 QY 2458 LysLeuSerGluGluLysMetGlnGlnGlnLysAspArgGluAlaSerThrIleGlnLeu 2477
 Db 6736 AAGCTTAAATATGATCAGACAGAAATATGATAGATTTCTCAAGTAATCTCTTT 6795
 QY 2478 ThrLysArgLeuLysAlaValAlaValGlnSerLysIle-----GlnArgGluIleThr 2494
 Db 6796 AATAACAAATTAATTCATATGATATGATATCAACAGATTTGAGAAAGAGCTTACC 6855
 QY 2495 ValTyrLeuAsnGlnPheGluAlaLysLeuGlnGlnLysGlnGlnLysGluLeu 2514
 Db 6856 ATA--TCCAAAGAGTGGGACAGACAGCTGAATCACTGAAAGAGAAATGAAAACTA 6912
 QY 2515 MetArgArgMetGlnHisHisGlyProSerAlaSerValMetGluGluGlnAspAlaArg 2534
 Db 6913 TTTAAAACTACCAAAACATTTGAAGCTTCTGTCATCTGTGCCACAGTTAAAT----- 6966
 QY 2535 LeuLeuGlyIleLeuLysThrValGlnAspGlu-----SerLysLys 2548
 Db 6967 -----CCTACCAACACAAATAAGAAATCCTCATGTTACTCAAGAGCT 7011
 QY 2549 LeuGln-----SerArgIleLysMetLeuGlnGlnGlnLysAsnLeuValLysAsp 2565
 Db 7012 ACACAGTTAACCAAGAGAAATATTCGAGAGCTGGAATATTCAGTCATGAGAGCTAAAGAA 7071
 QY 2566 AspAlaMetHisLysGlyGluLysValAlaIleLeuGlnAspLysLeu--LeuSerArg 2584
 Db 7072 AGTCTATGATGAAGAAAGCAAGATTAATAAGATGCAAGAAAGATTCAGTACTTAAT 7131
 QY 2585 AsnAlaGluAlaGluLeuAsnAlaMetGlnValLysLeuThrLysGlnAspAsnLeu 2604
 Db 7132 GACATTAATGCAAAATCT-----CAAGCCAAAGTTCATGATCAAAATTAATTCGCTT 7182
 QY 2605 GlnAlaAlaMetLysGluIleGlnLysAsnLeuGlnLysMetValAlaLysGlyAlaValPro 2634
 Db 7183 GAAAAAACAAGAGACAAATTCAGTACTCAGACAAAGATTCGTTTGAAGGCTTAAGCA 7242
 QY 2625 TyrLysGluGlnIleAspAsnLeuLysThrLysValValLysIleGluMetGluLysIle 2644
 Db 7243 TATTAAGAAAGAAATTAAGAGATCTCAAAATGAAGCTTTGAAATTAAGACTTGAAGAAATG 7302
 QY 2645 LysTyrSerLysAlaThrAspGlnGluIleAlaTyrLeuLysSerCysLeuGluAspLys 2664

QY 239 ILeaAsnArgSerLeuPheIleLeuGlnValIleLysLysLeuSerAspGluAla 278
 Db 862 ATAAATCGAAGCTTATTATTATTTGGACAAGATGATCAAGAACTTAGTCAGACAAGT 921
 QY 279 GLyGlyPheIleAsnTyrArgAspSerLysLeuThrArgIleLeuGlnAsnSerLeuGly 298
 Db 922 GGTGGTTTCATAAATATATACAGATAGCAAGTAAACAGAAATCTTCAGATTCCTGGGA 981
 QY 299 GLyAsnAlaLysThrValIleIleCysThrIleThrProValSerPheAspGluThrLeu 318
 Db 982 GGAATTCGAAGACACGATATTATCTGCACAAATTACTCCAGTATCTTTGATCAAACTCTT 1041
 QY 319 SerThrLeuGlnPheAlaSerThrAlaLysHisValArgAsnThrProHisValAsnGln 338
 Db 1042 ACTGCTTCACAGTTGCCAGTACCTCTAAATATATGAGAATATCCCTTATGTATGAG 1101
 QY 339 ValLeuAspAspGluAlaLeuLeuLysArgTyrArgGlyGluIleLeuAspLeuLysLys 358
 Db 1102 GTATCAACTGATGAAGCTCTCTGAAAAGTATAGAAAAGAAATATGATCTTAAAAA 1161
 QY 359 GlnLeuGlnAsnLeuGlnSerSerSerGluThrLysAlaGlnAlaMetAlaLysGlnGlu 378
 Db 1162 CAATTAGAG-----GAGGTTCTTTAGAGACGCGGCTCAGGCAATGGAAGAAACCA 1215
 QY 379 HisThrGlnLeuLeuAlaGluIleLysGlnLeuHisLysGluArgGluAspArgIleTrp 398
 Db 1216 TTGGCCCAACTTTTGCAGAAAAGATTTGCTCAGAAAAGTACAAATGAGAAAATTTGA 1275
 QY 399 HisLeuThrAsnIleValValAlaSerSerGlnGluSer-----GlnGlnAspGlnArgVal 417
 Db 1276 AACTTAACGCGATCGTGTGCTCTCTCTCCCTCAGCTTGAACAGAAATTAAGGCT 1335
 QY 418 LysArgLysArgArgValThrTyrAlaProGlyLysIleGlnAsnSerLeuHisLysLys 437
 Db 1336 AAAAGAAAACGACAGTACTTGCTGCTTGCGCAAAATTAACAAATGACAGAACTCAAC 1395
 QY 438 GlnValSerAspPheAspMetLeuSerArgLeuProGlnAsnPheSerLysLysAlaLys 457
 Db 1396 TATGCGATCATTTAAT-----ATACCAACAATATATACAAACAACACAT 1443
 QY 458 PheSerAspMetProSerPheProGluIleAspAspSerValCysThrGluPheSerAsp 477
 Db 1444 AAGCTTCTATTAATTTATTACGAGAAATGATGAATCTGTCTGTCAGAGCTGATGTT 1503
 QY 478 PheAspAspAlaLeuSerMetMetAspSerAsnGlyIleAspAlaGluThrAsnLeuAla 497
 Db 1504 TTCAGTACACTCTGTGATACATTAAGT-----GACATGCAATGCAATCCAGCA 1551
 QY 498 SerLysValIleHisArgGluLysThrSerLeuHisGlnSerMetIleAspPheGlyGln 517
 Db 1552 ACAAGCTACTAAATCAGAG----- 1572
 QY 518 IleSerAspSerValGlnPheHisAspSerSerLysGlnAsnGlnLeuGlnIleTyrLeuPro 537
 Db 1572 ----- 1572
 QY 538 LysAspSerGlyAspMetAlaGluCysArgLysAlaSerPheGluLysGluIleThrSer 557
 Db 1572 ----- 1572
 QY 558 LeuGlnGlnGlnLeuGlnSerLysGlnGlnGluLysGlnLeuValGlnSerPheGlu 577
 Db 1572 ----- 1572
 QY 578 LeuLysIleAlaGlnLeuGlnGluGlnLeuSerValLysAlaLysAsnLeuGluMetVal 597
 Db 1573 ----- 1587
 QY 598 ThrAsnSerArgGluHisSerIleAsnAlaGluValGlnThrAspValGluLysGluVal 617
 Db 1588 TTGAAGTCA----- 1596
 QY 618 ValArgLysGluMetSerValLeuGlnLysAspSerGlyTyrAsnAlaSerAsnSerAspLeu 637

Db 1596 ----- 1596
 QY 638 GlnAspSerSerValAspGlyLysArgLeuSerSerSerHisAspGluCysIleGluHis 657
 Db 1596 ----- 1596
 QY 658 ArgLysMetLeuGluGlnLysIleValAspLeuGluGluPheIleGluAsnLeuAsnLys 677
 Db 1596 ----- 1596
 QY 678 LysSerGluAsnAspLysGlnLysSerSerGlnAsnAspMetCysIleGlnLeu 697
 Db 1596 ----- 1596
 QY 698 CysGluAlaIleMetAlaGluLysAlaAsnAlaLeuGlnGluLeuAlaLeuMetArgAsp 717
 Db 1597 ----- 1605
 QY 718 AsnPheAspAsnIleIleLeuGlnLysGlnGluThrLeuLysArgGluIleAlaAspLeuGln 737
 Db 1606 GACTATGATTAATCTGGTATTACACTATGACAACACTACGACGAAAAGACAAATGCA 1665
 QY 738 ArgSerLeuLysGluAsnGlnGluThrAsnGlnPheGluIleLeuGlnLysGlnThrGln 757
 Db 1666 TTGAATTTAAAGAAAAGAAATGATTTGATGAAATTTGAGAGCTTGAAGAAAGAAACTAA 1725
 QY 758 LysGlnHisGlnAlaGlnLeuIleHisGluIleGlySerLeuLysLysLeuValGluAsn 777
 Db 1726 AAAGATCAAGAGATCAAAATTAATCATGAAATTTGCAAAATTAATTAATTAATTAAGAT 1785
 QY 778 AlaGluMetTyrAsnGlnAsnLeuGlnLysAspLeuGluThrLysThrLysLeuLys 797
 Db 1786 CGAGAAGTATATATCAAGATCTTGAGAAATGACACTGCTCAAAAGTAGAGCTGCTTGA 1845
 QY 798 GlnGlnGluIleGlnLeuAlaGlnLeuArgLysArgAlaAspAsnLeuGlnLysLysVal 817
 Db 1846 GAAAAGGAAGACAGATTAAGAACTACAGAAATCATGAGACTCTCAAAAGCTAAGAAAT 1905
 QY 818 ArgAsnPheAspLeuSerValSerMetGlyAspSerGluLysLeuCysGlnLulIlePhe 837
 Db 1906 ATAAATATGAGACTTGTCACTCACTCATG-----GAAAGCATTTGAACACCCAAA 1953
 QY 838 GlnLeuLysGlnSerLeuSerAspAlaGluAlaValThrArgAspAlaGlnLysGluCys 857
 Db 1954 CAATGAAGACACACTGTTGATCTGGAAGTACGCTTGATGACCAAGAGACATCA 2013
 QY 858 SerPheLeuArgSerGluAsnLeuGlnLeuLysGluLysMetCysLysPheSerAsnTrp 877
 Db 2014 GCCTTCTTAGAAGTGAATCTGAGATTGAGAGGAAATGAAGAACTTCAACTCA 2073
 QY 878 TyrAsnGlnLysGluLysAlaAlaSerLeuPheGluLysGlnLeuGlnThrLulLysSer 897
 Db 2074 TACAAGCAAAATGGAATGATATTCAGTATATCAAGCCAAATTTGGAGCAAAA----- 2127
 QY 898 AsnTyrLysMetCysGlnAlaAspLeuGlnLysGlnLeuGlnSerAlaPheAsnGluIle 917
 Db 2128 -----AAGAATATGCAAGTTGATCTGGAGAAAGAAATTAATCATCTGCTTTAATAGATA 2181
 QY 918 AsnTyrLeuAsnGlyLeuLeuAlaGlyLysValProArgAspLeuLeuSerArgValGlu 937
 Db 2182 ACAAACTCACCTCCCTTATGATGCAAAAGTCCAAAAGATTTGCTGTAATTTGGA 2241
 QY 938 LeuGlnLysLysValSerGluPheSerLysGlnLeuGlnLysAlaLeuGlnGlnLysAsn 957
 Db 2242 TTGGAGAGAAAGATTAGATCTTCAGAAAGAAATTAATTAAGAAATGAAGAAATGA 2301
 QY 958 AlaLeuGlnAsnGluValThrCysLeuSerGluTyrLysPheLeuProAsnGluValGlu 977
 Db 2302 GCTTGGCGGGAAGAGTCAATTTGCTTCAAGATTAATCTTTACCTTCAAGATAGAA 2361
 QY 978 CysLeuLysAsnGlnIleSerLysAlaSerGlnGluIleMetLeuLeuGlnGlnGlu 997

OY	1658	MetIleGluSerLysSerSerLeuGluSerLeuGlnHisGluYnHisAspThrGluGln	1677
Db	4474	-----AAAGAAACAAATAAAAGAAATTTGATAGCTAAACACCTGGAAACCTGAAGAG	4521
OY	1678	GlnLeu-----LeuAlaLeuLysGlnGlnMetGlnValValThrGlnGluLys	1693
Db	4522	GAACTTAAAGTTGCTCACTTCTTCCCTGGAAAGAACAGACAGAAACCTATTATGACGTGA	4561
OY	1694	LysGluLeuGlnGlnThrHisGlnHisLeuThrAlaGluValAspHisLeuLysGluAsn	1713
Db	4582	GTGAATCTTTACAGAGAAAGAA-----ACTGAAATATCAACCAATTCAAAGACAG	4629
OY	1714	IleGlu---LeuGlyLeuAsnPhelysAsnGluAlaGlnGlnLysThrThrLysGluGln	1732
Db	4630	TTACAAACCAATCATATGATTAATTTACGAACCAAGATCCAGACAAATTTATGACAAAAGGAA	4689
OY	1733	CysLeuLeuAsnGlnValAsnLysGluLeuGlnGlnSerGlnHisArgLeuGlnCysGluIle	1752
Db	4680	CAATTT-----AATATAAACCAATTTAGTACGCTTCAGGAAAAGTGAAT---GAACTG	4740
OY	1753	GlnGluLeuMetLysSerLeuLysAspLysGluSerAlaLeuGluThrLeuLysGluSer	1772
Db	4741	AAACAATTCAGAGGAGCATGCCAAGACCAAGATGATGACGCTCAAAAGTAT-----	4791
OY	1773	GluGlnLysValIleAsnLeu-----AsnGlnGluMetGluMet	1785
Db	4792	GAAAGTAAATGCTCGCGATGACCAACAGACTTCAGAAAGTCAAGAAATGCAATTT	4851
OY	1786	ValMetLeuGlnMetGlnGluLeuLysAsnSerGlnArgThrValIleAlaGlnAsp	1805
Db	4852	ATGATTTAAGGAAAGGAAAGGAAATGAAAGAGATACAGAGGCCCTTCAGATGACAGAGAC	4911
OY	1806	GlnLeuGlnAspLysLeuArgGluSerValGluMetSerIleGluThrGlnAspLysLeu	1825
Db	4912	CAACTGAAGAAACACTAAAGAAATTTGATGCTAAATGAAAGATCTCAAGAA-----	4965
OY	1826	ArgLysAlaGlnGlnAlaLeuGlnGlnGlnLysAspLysValGlnGluLeuThrSerGln	1845
Db	4966	AAACAAATATCACTTTCTTAGATGACAGCTGTCAATGACACTCAAGAAATATGTGAA	5025
OY	1846	IleSerValLeuGlnGlnLysIleSerLeu-----LeuGlnAsn-----	1858
Db	5026	ATACACACTTGAAGAGGACATTTGAGACCCGAAAGCTTAACCTGGAAGAAATGAAACG	5085
OY	1859	-----GlnMetLeuLys---AsnValAlaThrValLysGluThrLeu	1871
Db	5086	GAGAAATTAAGTTGACTGCATGACATACATGAAAAACCTTGAAGAAATGACATCTGTAACA	5145
OY	1872	SerGluArgAspLysLeuAsnGlnSerLysGlnHisLeuPheSerGluIleGluThrLeu	1891
Db	5146	AAAGAAAGATGACCTTAGAGAGTGTGGAGAGACTCTCAAAAGTGAAGAGACACAGCTC	5205
OY	1892	SerLeuSerLeuLysGluLys---GluPheAlaLeuGlnGlnAlaGluLysAspLys---	1909
Db	5206	AAGGAAACCTTAGAGAAACCTATTAAGACCTTGAAGAAACCAAGAGACCTAAATAAT	5265
OY	1910	-----AlaAspAlaIleArgLysThrIleAspIleThrGluLysIleSerAsn	1925
Db	5266	GTTCACATGCATCGAAGAGGACCAACAGAAATATTGAT-----AAACTAAGACGG	5316
OY	1926	IleGlnGluGlnLeuLeuGlnGlnAlaThrAsnLeuLysGluThrLeu---TyrGluArg	1944
Db	5317	ATTGTTTCAGAGAAACCAATGCAATATCAATTTATGCAAAAGAGACTTGTGAACACTCAAT	5376
OY	1945	GluSerLeu-----IleGlnCysLysGlnGlnLeuAlaLeuAsnThrGlnHis	1960
Db	5377	GATCCCTTAAACACACAGATCTGAAATATACAAAGAGGAACCTAAGATTTGCTACATGCAT	5436
OY	1961	LeuArgGluThrLeuLysSerLysAspLeuAlaLeuGlyLysMetGlnGlnArgAsp	1980
Db	5437	CTGAAAGACGACGAGAAACTATTTGCAAACTCAGAGGAATTTGTTCTGGAAGAACAGAT	5496
OY	1981	GlnAlaIleAsnLysValIleAlaLeuThrGlnLysMetSerSerLeuGlnGlnIle	2000

Db	5497	AAACTATCAAAATATGC	AAAAAGATTTCAGAAATTC	CAATGCTAAATTC	CAAGAAAGATT	5556
Qy	2001	AsnGlnAenValIThrThrLeu	LysGluGlyGluGlySer	ThrPhe	ThrLeuGln	2020
Db	5557	CAAGAA	-----CTTAAGCCAAATGAACATCA	CTTATACGT	TAATAAAACAT	5604
Qy	2021	ArgProSer	LysGlnGlnSerSerSerGln	MetGluGluLeuArg	GluSerLeuLysThr	2040
Db	5605	GTCATAGAGACACAGAA	AAAAAGTCGTGAATG	GACCACTAAAGCAATTA	AAAGAC	5664
Qy	2041	LysAspLeu	GlnLeuGlnGluAlaGluGly	LeuLeuSerGluAla	ThrAsnGluLeuLys	2060
Db	5665	CAAAAGCTTA	CTGAGCTAAATTA	-----	-----GAAATTAAG	5697
Qy	2061	AsnLeuThrAlaLys	LeuSerSerLeuGluGluGluLeu	GlnAsnAlaSer	ThrLeu	2080
Db	5698	AAATTA	-----AAATTCGCTCAAGAACTT	CATGAAACCTG	GAAGAAATG	5742
Qy	2081	AsnGluAlaVal	SerGluArgGluAsnLeuArg	HisSerLysGlnGlnLeuVal	SerGlu	2100
Db	5743	AAATCTGTAATGAA	AAAGAGATTAATCTAAAG	AGAGTGAAGACACATCA	AACTGGAG	5802
Qy	2101	LeuGlnGlnLeuSerLeuThr	Leu-----LysSerArgAsp	HisAlaPhe	LeuGln	2117
Db	5803	AGAGCCAACTCAAG	GAAGAAAGCTCTGCAAGAAAC	CAAGCTTAGAGATCTG	GAATACACAG	5862
Qy	2118	-----	-----SerLysArgGluLys	-----AspGluAlaVal	-----	2126
Db	5863	GAACATAAAATGCTCT	GATGCTATCAAAAGACACAAAGAA	AACTGTGATTA	ACTTGA	5922
Qy	2127	AsnLysThrLeuAlaSerLeuAla	GluGluGluLeuThrLysGluGlu	MetSerAspGluPhe	-----	2146
Db	5923	GAATAAAATTCAGAA	AAAGACAAATTCAAATTCAG	CAATTCAAAGAGATTTAGATTA	---	5979
Qy	2147	ArgAspSerLysGluSerLeuGlnGln	GlnSerSerHisLeuSerGlu	LeuLeuCysThr	-----	2166
Db	5980	-----TCAAAAGATGA	ATTATTCAGAAAAAGATCC	CAAGAACTTCAAG	-----	6018
Qy	2167	TyrLysThrGlnLeuGln	MetLeuLysGlnLysGluAsp	PheAsnAsnLysLeuAla	-----	2186
Db	6019	-----AAAAAGAACTTCA	AGCTTGAAG-----GTAAAGAA	GATGTCATTAATGAGT	---CAT	6069
Qy	2187	GluLysValLysGluVal	AspGluLeu-----	-----	-----	2195
Db	6070	AAAAAAATTAATGAAT	GAAGAAAGCTTGAAGACAAATTT	GACCAAACTATG	TCAGAG	6129
Qy	2196	-----	-----LeuGlnHisLeuSerSerLeu	LysGlnGlnLeuAspGlnLeuGln	-----	2210
Db	6130	TGTGAGATGATTA	CTTCAGGTTGAGCTAAGAAAC	TTCAATGAAAGCTTGAAGAA	ATAAGA	6189
Qy	2211	MetGluLeuArg	---AsnGluLysLeuAlaArgAsn	TyrGluLeuCysGluLysMetSer	Phe	2229
Db	6190	ATTGAGCTAAAGAA	GAGATGAGCTAAGAGGAGAT	AAAAATCTCTCAAAATG	AGAA---	6246
Qy	2230	MetGluLysGluLeuSerVal	LeuArgLeuMet-----Gln	-----	-----	2241
Db	6247	AGGACCAATTCATAG	CAACCTTAAGGAAATGATAG	CTAGAGACCGACAGAC	CAACCA	6306
Qy	2242	AsnGluProGlnGlnGlu	-----GluAspAspValAlaGlu	ArgMetAspPheLeuGlu	-----	2259
Db	6307	GTAAGAACTGAAAA	AAAGCTTAAGATGAGCAAC	ACACCTTAAGAAAGCTG	AGAA	6366
Qy	2260	SerArgAsnGlnGluGlnGln	GluMetGluLysLeuSerAlaVal	TyrSerGluGln	-----	2279
Db	6367	GAAGAAGCTCTAG	AAATTAAGAGCTTTGAAGAG	TACTGACAGATG---GATGATCAT	-----	6423
Qy	2280	HisThrLeuLeuSerSerLeuSer	GlnGluGlnGluGluGluGluGlu	HisAlaHisLysHis	-----	2299
Db	6424	TATGAGTGCTTGAAT	TAGATTGCTCTTGAC	TTGAGAAAGAAATGGAAT	TTCCACAGA---	6480
Qy	2300	CysMetLeuAsnLysGluSerLeu	SerSerThrLeuSerArgSerPhe	GlySerLeu	-----	2319

Db 6480 ----- 6480
 QY 2320 GlnThrGlnHisValLysLeuAsnThrGlnLeuGlnThrLeuLeuAsnLysPheLysVal 2339
 Db 6481 -----ATCAGTGAAGAAACGCAAGTAT 6501
 QY 2340 Val-----TyrArgThrAlaAlaValLysGlnAspHisSerLeuIleLysAspTyrGlu 2357
 Db 6502 GTCTTAGCTATGTTGAGTGAAGTGAAGAAACCAACATGATGATCATTAATTAATTGAA 6561
 QY 2358 LysAspLeuAlaIleGlnGlnLysArgHisAspGlnLeuArgLeuGlnLeuGlnCysLeu 2377
 Db 6562 ATGGATTTTGTGATGAGTGAAGTGAAGAAACCAACATGATGATCATTAATTAATTGAA 6621
 QY 2378 GlnGlnHisGlnLysArgLysTyrSerAspSerAlaSerGlnLysLysPheCysGlnIle 2397
 Db 6622 CAACAAGAT-----TGTGATGTACATCCAGACATTAAGGATCTCAAAATTG 6669
 QY 2398 GlnPheLeuAsnGlnLeuLeuPheLysLysAlaAsnIleIleGlnSerValGlnAspAsp 2417
 Db 6670 AACCAAGATATGATCTA-----CAATGTGAGAAATTCCTCAAGAT 6711
 QY 2418 PheSerGlnValGlnValPheLeuAsnGlnValGlySerThrLeuGlnGlnLeuGln 2437
 Db 6712 TTCTCAGAAAGTGAG-----TTCCTCAGCATTAAGACTGAATTCCTCAACAGTACTAAGT 6765
 QY 2438 HisLysLysGlnPheMetGlnTyrLeuGlnLysPheGlnLysPheHisValAspAlaLys 2457
 Db 6766 AATAGGAAGAAATGACACAGCTTTTGGAGAGTGGTTAATCTCGTTTGTGATATGAA 6825
 QY 2458 LysLeuSerGlnLysMetGlnGlnGlnLysAsnArgIleAlaSerThrIleGlnLeuLeu 2477
 Db 6826 AACCTTAAATATGCTCCGAAAGAAATGATGATGATGCTCAAGTAACTCTT 6885
 QY 2478 ThrLysArgLeuLysAlaValAlaGlnSerLysIle-----GlnArgGlnIleThr 2494
 Db 6886 AATAACAGATTAATTCCTCAATGAATGAATCAACAGAGTTTGGAGAAAGAGTCTACC 6945
 QY 2495 ValTyrLeuAsnGlnPheGlnAlaLysLeuGlnLysLysGlnGlnAsnLysGlnLeu 2514
 Db 6946 ATA-----TCCAAAGAGTGGAGACGACCTGAATCACTGAAGAGAAATGAAGAACTA 7002
 QY 2515 MetArgArgMetGlnHisHisGlyProSerAlaSerValMetGlnGlnGlnAsnAlaArg 2534
 Db 7003 TTTTAAATCAACCAACATTTGAAGACTTCCTGTGGCATCTGGTCCAGGTTAAT----- 7056
 QY 2535 LeuLeuGlnIleLeuLysThrValGlnAspGln-----SerLysLys 2548
 Db 7057 -----CCTACACACAAAGACAAATGAATCCTCATGTTCATCATCAAGAGCT 7101
 QY 2549 LeuGln-----SerArgIleLysMetLeuGlnAsnGlnLeuAsnLeuValLysAsp 2565
 Db 7102 ACACAGATTAAACACAGAGAAATTCGAGAGCTGAAATTCATCTGAAAGCTGAAGAA 7161
 QY 2566 AspAlaMetHisLysGlnLysValAlaIleLeuGlnAspLysLeu-----LeuSerArg 2584
 Db 7162 AGTGCATGATGATGAAGAAAGCAAGATTAATGAAGTGCAGAAAGCTTGAGGACTAAT 7221
 QY 2585 AsnAlaGlnLysGlnLysLeuAsnAlaMetGlnValLysLeuThrLysLysGlnAspAsnLeu 2604
 Db 7222 GACATATAGCAAACTT-----CAAGCCAAAGTTTCATGAATCAATTAATGCTTT 7272
 QY 2605 GlnAlaAlaMetLysGlnIleGlnAsnLeuGlnLysMetValAlaLysGlnAlaValPro 2624
 Db 7273 GAAAAAAGCAAGACATTAAGTAACTTCAGACAAAGTCTTTAGAGAGTAAAGCCA 7332
 QY 2625 TyrLysGlnGlnLysAspAsnLeuLysThrLysValValLysIleGlnMetGlnLysIle 2644
 Db 7333 TTTTAAAGAAATTTGATGATCTCAAAATGAAGCTTGTGAAATTAACCTAGAGAAAGT 7392
 QY 2645 LysTyrSerLysAlaThrAspGlnGlnIleAlaTyrLeuLysSerCysLeuGlnAspLys 2664
 Db 7393 AAAAAATGCCAAGAAATTTGAAAAAGAAATCAGTCAACAAAGCCACTGTGAATATCA 7452

QY 2665 GlnGlnLysLeuArgArgLeuLysGlnGlnLysLeuArgArgAlaIleAspAsnAspThr 2684
 Db 7453 AAGAGATTTATAGGCGATTTGAGAGAAATCTCGAAGAGAGTCAACAGGCCCAAGATAC 7512
 QY 2685 ThrValCysValProLysAspTyrGlnLysAlaSerThrPheProValThrCysGly 2704
 Db 7513 TCAGTGTATTCAGAACATCTGATCCACGCTTCAATTAACCTTAACCTTGTGGAGT 7572
 QY 2705 GlySerGlyIleValGlnSerThrAlaMetLeuValLeuGlnSerGlnLysAlaLeu 2724
 Db 7573 GGCAGCGCATTTGATACAAACAAAGAGCTTTATTTGAAAGTGAACATTAAGGCTA 7632
 QY 2725 GlnArgLeuLeuSerHisTyrLysLysTyrHisHisSerAlaArgThrMetSerSer 2744
 Db 7633 GAAAAAGAAATTTCTAGTTAAAGCAGCAAAATGACAGCTA----- 7674
 QY 2745 SerGlnAspArgLysThrLysAlaLysSerAspAlaHisSerSerHisThrGlySer 2764
 Db 7675 -----ATMAAACAAAGAAATGATTTTAAGCAATATCAGATCTT 7716
 QY 2765 SerHisArgLysSerProHisLysThrGlnThrTyrArg-----HisGlyProVal 2781
 Db 7717 TCCAAATGAGTCAAAACTTGGAGAAAGAACCTTAAAGAGAGGCTCAAAACAAAGTA 7776
 QY 2782 ThrProGlnArgSerGlnMetProSerLeuHisLeuGlnLysSerProLysSerGlnSer 2801
 Db 7777 ACTGTGAG-----AATCTCCAAAG----- 7797
 QY 2802 SerThrLysArgValAlaSerProAsnArgSerGlnIleTyrSerGlnLeuValMetSer 2821
 Db 7798 -----TCT 7800
 QY 2822 ProGlnLysThrGlnMetHisLysHisIleLeuSerProSerLysValGlnLeuHisLys 2841
 Db 7801 CTTAAAGTGCATGAGACA-----GCTTCTAA 7827
 QY 2842 LysArgAlaLeuSerProAsnArgSerGlnMetProThrGlnHisValIleSerProGln 2861
 Db 7828 AAGAAACAAATTTACACC----- 7845
 QY 2862 LysThrGlnLeuHisLysAsnLeuThrGlnSerThrLeuPheAspAsnLeuSerPro 2881
 Db 7846 -----TCTCA 7851
 QY 2882 CysLysGlnGlnLysValGlnGlnLeu-----AsnSerProLysGlnLysLeuPhe 2899
 Db 7852 TGCAGAGAACGAAATTTACAGATCCTGTGCCAAGAAATCACCAAAATCTTTT 7911
 QY 2900 AspValLysSerLysSerMetProTyr--CysProSerGlnPhePheAspAsnSerLys 2918
 Db 7912 GATAGCCGATCAAAAGCTTTTACATCATCCATTCAGAGTGGCTATTTTGTATCTCAAG 7971
 QY 2919 LeuGlnAspPheSerGlnLeuAsnThrAlaGlnSerAsnAspLysSerGlnAlaGln 2938
 Db 7972 TTAGGCTTTTGCAGAGGTGCAAAATGCAAGAGAGAGTGTGATTCCTCAGCAGGT 8031
 QY 2939 TrpTyrGlnAlaLysLysGlnThrAlaProGlnCysLysThr 2953
 Db 8032 CCTTGGCAGCCCTCTCAGGCAAGATGTCCGTGATGCAAACT 8076
 RESULT 3
 US-09-960-253-145
 ; Sequence 145, Application US/09960253
 ; Patent No. US20020123619A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Benson, Darin R.
 ; APPLICANT: Mohamath, Raodoh
 ; APPLICANT: Lodes, Michael J.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121.556
 ; CURRENT APPLICATION NUMBER: US/09/960,253


```

1 CURRENT FILING DATE: 2001-09-20
2 NUMBER OF SEQ ID NOS: 187
3 SOFTWARE: fastseq for Windows Version 4.0
4 SEQ ID NO 145
5 LENGTH: 10300
6 TYPE: DNA
7 ORGANISM: Homo sapiens
8 US-09-960-253-145

```

Assignment Scores:	
Pred. No.:	1,01e-64
Score:	1153.00
Percent Similarity:	39.53%
Best local Similarity:	20.13%
Query Match:	7.81%
DB:	10
	Gaps:
	138
	Indels:
	858
	Mismatches:
	1170
	Conservative:
	651
	Matches:
	675
	Length:
	10300

US-09-150-867-1 (1-2954) x US-09-960-253-145 (1-10300)

OY	110	GlValPheLysIleIleGlnGluIleProAsnArgIlePheLeuAlaValSerTyr	129
Dd	559	GAAATGAAGAAAAGATAAACAATTAGCTCCACGAGAGAGGAACTATTC-----	606
OY	130	MetGluIleTyraSncGluThrValLysAspLeuCysAspAsparGadgLyIsPro	149
Dd	607	-----AACACTTTGCCAGCCCCAGCTTCTCAGGCACACGGCAACAACCT	651
OY	150	LeuGIuileArgGluAspPheAsnArgAsnValTyrValAlaAspLeuthrGluGluLeu	169
Dd	652	GCACACGAGTCTCACAGAGATG-----GAAAGAATTG	681
OY	170	ValMetAlaProGluHisValIleGlnTrpIleLysLysGlyGluLysAsnAlaGlnSTyr	189
Dd	682	GTAAATGATGAAGCACAGCTC-----CAGAGAAGGAAGTAATTCATT	723
OY	190	GlyGluThrLysMetAlaAspHisSerArgSerHisThrIlePheArgMetIleVal	209
Dd	724	AGCATTTACAGGCCCGCCTCAGCGACACACAGGCAGACGAACTGCACACAGAGTGCTC	783
OY	210	GluSerArgAspArgAsnAspProThraSnSerGluAsnCysAspGlyAlaValMetVal	229
Dd	784	CGACAGAAAGATGCCCTTTGAACAACAGATTCGTTCATGAAGATGAGCTT-----	837
OY	230	SerHisLeuAsnLeuValAspLeuAlaGlySerGluArgAlaSerGlnThrGlyAlaGlu	249
Dd	838	-----CTTCAGTTAGTAAACCACGACGATGTGGAAACAGAGATGCAA-----	879
OY	250	GlyValArgLeuLysGluGlyCysAsnIleAsnArgSerLeuPheIleuGlyGlnVal	269
Dd	880	-----CAGAAATTGAGGGTGCTGCAGAAAGAG	906
OY	270	IleLysLysLeuSerAspGlyGlnAlaGlyLysPheIleAsnTyrArgAspSerTysLeu	289
Dd	907	CTTGAGGAAACCGAAGAAATCCTTGGTGGGC-----CGTGGTCAGGTC	948
OY	290	ThrArgIleLeuGlnAsnSerLeuGlyGlyAsnAlaLysThrValIleIleCysThrIle	309
Dd	949	GTTGACTCTGCTGACACAGGAGCTGACTGCTGTGACGACAGAA-----	990
OY	310	ThrProValSerPheAspGluThrIleSerThrLeuGlnPheAlaSerThrAlaLysHis	329
Dd	991	-----AACCGAATTCCTCTCCACAGCTTACAGCAGATGCAAGCTGGAGCAAT	1033
OY	330	-----ValArgAsnThrProHisValAsnGluValAlaLeuAspAspGluAlaLeuLys	347
Dd	1036	AATACTTTGAGGAACACTGTGGGA--ACGAAAGAGAGAGATGCCAAGATTTACTGGGA	109
OY	348	ArgTyrArgLysGluIleLeuAspLeuLysGlyGlnLeuGluLysLeuGlnLeuSerSer	367
Dd	1093	AAGATGGAACCTGAAGGGGAGAGAAAAATTATCTTCATATCTGCCAGAAAGAAATG	115
OY	368	-----GluThrLysAlaGlnAlaMetAlaLysGluGlnHisThr	380

Dd	1153	CATGATCTTTTAAAGAAKAGSTTGAGCAAGACGGCCAAAGGCCAGGCGTAGACTAAG-----	1208
Oy	391	GlnLeuLeuAlaGluIleLysGlnLeuHisLys--GluArgGlnAspArgIleTrpHis ::: :::	399
Dd	1207	TTCGGGTATAGTGGTTTTGGAGCAGAAGACAAGAAAGCAAGAAATGGAAAGAACSTCTCAT	1266
Oy	400	LeuThrAsnIleValValAlaSerSerGln ::: :::::	409
Dd	1267	ATTGTGAAGCTCTCAAAGAGCTGCACAAAGAGCTGCAGTGGCTGTGATGCTTAAGAT	1336
Oy	410	-----GlnSerGlnGlnAspArgValIleValArgVal ::: :::	423
Dd	1327	CAAAATTCAAAAGCTCTCCCAAGATAGATGAAAACAGAGCAATTCAGTCAAGCCAGCAAT	1386
Oy	424	ThrTrpAlaProGlyValIleGlnAsnSerLeuHisAlaSerGlyValSerAspPheAsp ::: :::	443
Dd	1387	CAGAACTGGAAAGATCCAGCTSCAGCAAAAA-----TCCAAAGAAATGAGCCAAATT---	1437
Oy	444	MetLeuSerAlaGlyProGlyAsnPheSerLysValValLysPheSerAspMetProSer ::::	463
Dd	1438	--STAATTAACAGTCCC-----TTCCAAACAATGAACAAACAGCAATCAGACTCT	1489
Oy	464	PheProGluIle-----AspAspSerValLysThr-- :::::	473
Dd	1486	TTCCCAAGATGTTTAAATGAGGCSACASACAGGCACTAGAGAGATATTTCTTTCTTGG	1545
Oy	474	-----GlnPheSerAspPheAspAspAla-----LeuSerMetMetAsp :::: :::::	486
Dd	1546	CAGAAAGAGAGTGGTAAAGATAAGAAATGAACAAAGAGAGCTGGCTGTTAGTTCTATAG	1605
Oy	487	SerAsnGlyIleAspAlaGlu---TrpAsnLeuAlaSerLysValThr----- ::: ::: :::	501
Dd	1606	CTGGAGGAGCTGAAGAGCTGAGAAATGAACAACTCTCTCAATTAAGTCTCTCTAGAGCT	1665
Oy	502	-----HisArgGlyLysThrSerLeuHisGlnSerMetGln :::: ::: :::	514
Dd	1666	CAGATTAAGACTGGGAGGAGCACAGACAGAAAGTAGTAGAGATC-----AGCATTTTAT	1719
Oy	515	Phe-----GlyGlnIleSerAspSerValGln ::: ::::	523
Dd	1720	ATTGGCAACAGAGAGACTCTGTGCTGAGAAAGTGGACAA-----CATCTCTAGAA	1773
Oy	524	PheHisAspSerSerLysGlnAsnGlnLeuGlnIleTrpLeuProLysAspSerGlyAspMet :::: ::::	543
Dd	1774	AACACATTTTCTCGAAACACATTAAGATTTCTCAGTTT-----TTG	1815
Oy	544	AlaGlnCysArgLysAlaSerPheGlnLysGlnIleThrSerLeuGlnGlnIleGln :::: :::	563
Dd	1816	TTGGAAATGAAGAAAGCT-----CAAGAGAAATATGCATTTCTTAATATTCAGGCTCAG	1869
Oy	564	SerLys-----GlnGlnGlnLysLysGlnLeuValGlnSerPheGlnLeuLysIleAla :::: :::	581
Dd	1870	GGAAAAAGGCGCTAGAGAGCAAGATCATAGGCTCTTGACCAGAAAGAAATGAACAGATG	1929
Oy	582	GlnLeuGlnGlnGlnLeuSerValLysAlaLys-----Asn :::: ::::	593
Dd	1930	GAGGCTGAGGGAATAGCTCCAATTAATGAATGAATTTCTTGAAGATCAGCGCAAGAT	1989
Oy	594	LeuGlnMetValThrAsnSerArgGlnHisSerIleAsnAla-----GlnValGlnThr :::: ::::	611
Dd	1990	TTTCCCTTAAGGCCAAAT---GAAGAGAGCAGCTTCCAGCACTGAAACAAACAGAGCG	2046
Oy	612	AspValGlnLysGlnValValAlaArgLysGlnMetSerValLeuGlyAspSerGlyTrpAsn :::: :::	631
Dd	2047	AGCACTGAACATCAAAATAGACATCTGAGGAATATCTTTAAATAGATGCTGAGATA---	2103
Oy	632	AlaSerAsnSerAspLeuGlnAspSerSerValAspGlyLysArg----- ::: :::: ::::	646
Dd	2104	-----GAATTGAATCAACAAACAGAGATGGTGAATAATCCCTTCTGCTGTA	2151
Oy	647	-----LeuSerSerSerHisAspGlnLysIleGlu----- ::: :::: ::::	656
Dd	2152	CCAGATTTTGGCTAGTCTCTCAAGATCAAGAGATCAAGAAAGGTTAAAAAGTCAATTTTGGAG	2211

QY 657 -----HisArgLysMetLeuGlnLys---- 664
 Db 2212 CTCGAGCTAAACTTTCATTAACACAGAAATCTATGAGAAAATTTAGATAGAAAAGCT 2271
 QY 665 -----LLevalAspLeuGlnGluPheIleGluAsnLeuAsnLysSerGluAsnAsp 682
 Db 2272 AAGGAATTTAGCAACCTTAACCCAGTTGATTTGAGAGATTAAAGAAAATGCTGACAAACAAC 2331
 QY 683 LysGlnLys-----SerSerGluGlnAspPheMetGlnSerIleGlnLeuLys 698
 Db 2332 AGCAATGCAATTCACCTGCTTTTGTCTCAAGAGAAAGACAGCTTCTCTC---CAGGTGAAG 2388
 QY 699 GluAlaIleMetAlaGlnLysAlaAsnAla---LeuGlnGluLeuAlaLeu----- 714
 Db 2389 GAACTTAGCATGTGTAAACAGAAATTGAGGCTCAGTAAAGCAAACTGGAATGACCTTGCA 2448
 QY 715 -----MetArgAspAsnPheAspAsnIleIle 723
 Db 2449 GAGCAGAAAGGCAAGAAAGACTGATTATGAAAGCCAAACTGCCCATGACAACTGCTC 2508
 QY 724 LeuGluAsn-----GluThr 728
 Db 2509 ACTGAACAGATCCATAGTCTCAGCATAGACCAATCTAAAGATGTGAAAATTGAACTT 2568
 QY 729 LeuLysArgGluIleAlaAspLeuGln-----Arg 738
 Db 2569 TTACAGAAATGAACTGATGATGTGACGCTTCAGTTTCTGAGCAGAGTACCCTGATAGA 2628
 QY 739 SerLeuLysGluAsnGlnGluThrAsnGlnPheGluIleLeuGlnLysGlnThrGlnLys 758
 Db 2629 AGCTCGCAAAACCCAGCTCCAAATAAGAGAAAGTGAAGTGGCTTGAAGGGGCGCAAGCTGTA 2688
 QY 759 GluHisGlnAlaGlnLeuIleHisGlnIleGlnSerLeu-----LysLysLeuValGlu 776
 Db 2689 AGGATATCTCAACTAAAGTGAAGAACTGTCCACAGCTTTTCAAGAAAGCACTTGAA 2748
 QY 777 AsnAlaGluMetLysArgAlaAspAsnLeuGlnGlu-----AspLeuGlnThrLysThrLys 794
 Db 2749 ATAACAAAATATGATCAGCTCTTACTAGAGAAAAGAGATGTGAAACCCCTCCAAACA 2808
 QY 795 LeuLeuLysGlnGlnIleGlnLeuAlaGlu----- 805
 Db 2809 ACCATCGAGGAGAGATCAACAAGTGAACATCAGCTTATGATGACTGAGAAATG 2868
 QY 805 ----- 805
 Db 2869 GTTCAGCTTAATGAGAAAGTTTCTCTGGGGTTGAATTAAGACTCTTAAGAACAG 2928
 QY 806 -----LeuArgLysArgAlaAspAsnLeuGlnLysValArgAsnPheAspLeuSer 823
 Db 2929 CTAATATTATTCACAGCTGAGAGCAAAAAGAGAGAGGTGAGAGAAATATGAA 2988
 QY 824 ValSerMetLysAspSerGlnLysLeuGlnIlePheGlnLeuLysGlnSer--- 842
 Db 2989 GTTCTCTCTGCGC-----CTTAACAAAATATAT 3015
 QY 843 -----LeuSerAspAlaGlnAlaValThrArg---AspAlaGlnLysGlnLysSerPhe 859
 Db 3016 GATGAGATGTAGCCAGAGCAAAATAGTAAGAAAGAACTTCAGCATGAAATTTGACCTT 3075
 QY 860 LeuArgSerGluAsnLeuGlnLeuLysGlnLysMet----- 871
 Db 3076 CTGAGAAAGAAATATGAGCAGAGAAAGAAAGAACTCCAGCAGCTCTTATTAACAGAAAG 3135
 QY 872 -----GluAspThrSerAsnThrPyrAsnGlnLys 881
 Db 3136 GACCTTCTGCAAAAGCTCAGATGATTGAAAGAAATTAACCAACTTGAAGATGATCT 3195
 QY 882 GluLysAlaAlaSerLeuPheGlnLys-----GlnLeuGlnThrGlnLysSerAsn 898
 Db 3196 AAGAAAGAAATCCCACTCAGTGAAGCTGAGAGGGGAGAGAGTGAAGAAATTAAGAAAG 3255

QY 899 TyrLysLysMetGluAlaAspLeuGlnLysGlnLeuGlnSerAlaPheAsnGlnIleAsn 918
 Db 3256 AAGAAATCTCAGAA-----AAATGTGTACTTAAATGCGCAAGAAATGAA 3303
 QY 919 ---TyrLeuAsnGlnLysLeuAlaGlnLys-----ValProArgAsp 931
 Db 3304 ATTTATTTAAACAGACATATCTGAGAAAGAAAGTGAACATACAGCATATATAGAGAGAT 3363
 QY 932 LeuLeuSerArgValGlnLeuGlnLysLysValSerGlnPheSerLysGlnLeuGlnLys 951
 Db 3364 TTGGAGAAAGTGTGCGCAGCTGAGAGACCAATTCAGGCTGTGTCAAACAGATGAATCAG 3423
 QY 952 AlaLeuGlnGlnLysAsnAlaLeuGlnAsnGlnValThrCysLeuSerGluTyrLysPhe 971
 Db 3424 ACCTTGCAGATTAAACA----- 3441
 QY 972 LeuProAsnGlnValGlnCysLeuLysAsnGlnIleSerLys----- 985
 Db 3442 -----AACCAATATAGATTGCTCCAGCAGAAATCAGTGAAMAAACAAAGCAATTATCAG 3495
 QY 986 -----AlaSerGlnGluIleMetLeuLysGln 995
 Db 3496 AAGTTAATCACAAAGTAAACAGGATGCAAGTATGGGAGCTCCGTAGCATTTGTAAGGAA 3555
 QY 996 GluGlnGluHisSerAlaSerIleIleSerLysGlnGluIleIleMetGlnGlnIns 1015
 Db 3556 ACACTGTGATTAAGTCCACCTGTGACAGTAGTAGTAAACATCGAAACCAAGAACTAGAA 3615
 QY 1016 GluGlnIleLeuGlnIleThrAspGlnValThrHisThrGlnSerLysValGln----- 1033
 Db 3616 GMAAGATAGTGGCTGTGAAAAGAAAGAGCAACCTTCAAAAAGAGTACAGAAAGCC 3675
 QY 1034 -----GluThrGlnGlnGlnIleu----- 1040
 Db 3676 TTAACTCCCGCAAGCAATTTTAAAAAGCACAGAGAAAGAAACACTTCACGGAG 3735
 QY 1041 GluMetLysLysMetHisAspAspLeuPheGlnLysTyrIleArgAsnLysSerGlnAla 1060
 Db 3736 GACTTAAGCAACAGAAAGATGAC-----TATTAATGCTTGCAAGAAAGATT 3783
 QY 1061 GluAspLeuAsnArgGluMetGluAsnLeuLysGlnIleThrMetLysSerValGlnValLys 1080
 Db 3784 GATGAGCAAAAGCAAGGAAATGAAATATTGAGACCAAGTAAAGCACTCCAGATTCAA 3843
 QY 1081 IleAlaAspThrLysHisGlnLeuGlnGluThrIleArgAspLysGlnLeuLeuHis 1100
 Db 3844 GTAAGGAATCC----- 3855
 QY 1101 GluLysLysTyrPhePheGlnAlaMetGlnThrIlePheProIleThrProLeuSerAsp 1120
 Db 3856 -----ATAGACGGA 3864
 QY 1121 SerLeuProProSerLysLeuValGlnGlnLysAsnSerGlnAspProIleGlnLysAsnAsp 1140
 Db 3865 AAACCTCCCAAGCAGACACAGCAGGAAATGCTGTCTCCACTCCA-----GGT 3912
 QY 1141 TyrHisAsnLeuIleAlaLeuAlaThrGluArgAsnAsnIleMetValCysLeuGlnThr 1160
 Db 3913 TTAGAGAACACTTTATTAAGCCAGCAAGAACAGATCAACACCAACCTGTTTAAGAGTCC 3972
 QY 1161 -----GluArgAsnSerLeuLys-----Glu 1167
 Db 3973 AACTGTGCCCAAGACTGGCTTCTCATTTCTGAAGATGCGACGCTGTGAGGGCGGAAGT 4032
 QY 1168 GlnValIleAspLeuAsnThrGlnLeuGlnInsLeuGlnAlaGlnSerIleGln----- 1185
 Db 4033 TCGTTGGCCCAATTAAGGCCCAAGCTGAAAGGAAATAGAGCTGAAAGATAGAGTTAGAA 4092
 QY 1186 -----LysSerAspLeuGlnLysProLysGlnAsp-----Leu 1196
 Db 4093 TTGAAGTATGATTCTAACAACAGTGAAGTACTATAAAAATCAGAAAGAGATATTACAGTTA 4152
 QY 1197 GluGlnGlnGlnValLysLeuLeuLeuGlnMetGlnLeuLeuLysGlnIleHisLeuThrAsp 1216

Db 4153 CAACGACGATAAATTAACAGAGCTTAAAGATCGAGACTTAAGACAGTATCCCATGAA 4212
 Qy 1217 SerLInleuSerLleuGluLysLeuGlnLeuGlnValThrGluLysLeuGln 1236
 Db 4213 GCTAAAGTCCATGCCGAAAGGCTGCAG-----CAAGAAATTTGGAA 4251
 Qy 1237 ThrLeuGlnGluMetLysAsnIleThrIleGluArgAsnGluLeuGlnThrAsnPh 1256
 Db 4252 AGCAGCCACACACAAATTCGCGCTAGACACTTAAAG---GAATTCACCACTTAAGTG 4308
 Qy 1257 GluAspLeuLysAlaGlnHisAspSerLysGlnAspLeuSerGluAsnIleGluGln 1276
 Db 4309 GATGAAGCTGCAAAAATCATCAAGCAAAAAGAGACAGCTTACTACTCTTCGACAA 4368
 Qy 1277 SerIleGluThrGlnAspGluLeuArgAlaIleGlnGluLeuArgGluLysGln 1296
 Db 4369 CTTAGTGAAAGAACGACCTCTCACTTAATACGACAGAGATTAATGAAACAAAGCAT 4428
 Qy 1297 LeuValAspSerPheArgGlnGlnLeuLeuAspCysSerValGlyIleSerSerProAsn 1316
 Db 4429 TTAATTAAGGCTCGCATACACAGCTA-----GAATTCACCAAGCCAAAGAG 4473
 Qy 1317 HisAspAlaValAlaAsnGlnGluLysValSerLeuGluValAsnSerLeuGlnSer 1336
 Db 4474 CATGATGAGAGATTAAGACAGCTACAGGTGACACTTGTGAAATGAAACAAACAGAA 4533
 Qy 1337 GluMetLeuArgGlyLysAlaArgAspGluLeuGlnThrSerCysLys-----AlaLeu 1353
 Db 4534 GAGATTGAGAGAAAGAAAGTAAAGCAACAAACAAATACAAAGAAATGCAAGCTGCCCTT 4593
 Qy 1354 ValSerGluLeuGluLeuLeuArgAlaHisValLysSerValGluGlu----- 1370
 Db 4594 ATTTCCGAAAGAAAGAACACTTAAGAAAC---AAAGTCTCCAAAGAGAAATTGCTTTG 4650
 Qy 1371 -----AsnLeuGlu-----IleThrLysLysLeuAsnGluLysGluLysGluLeu 1386
 Db 4651 GCCAGAGTACCATTTGAACGCTCACCACAAAGCTCTGCGCAGATGTGAAAGCAAGTTCT 4710
 Qy 1387 GlyLysSerGluGluSerGluValLeuLysSerMetLeuGlnAsnLeuLysGluAspAsn 1406
 Db 4711 GCTCAAAATTAAGAAAGATACGCTCTTAAGAAAGTTAGCTTTCTTCAAGAAAGAA 4770
 Qy 1407 AsnLys-----LeuLysGluGlnAlaGluGluTyrSer 1417
 Db 4771 GACAAACTCATTAAGAAATGACAGCTCTTATTGAAATACAGAGCTCAGCAGCTCC 4830
 Qy 1418 SerLysGluAsnGlnPheSerLeuGluValPheSerGlySerGlnLysLeuValAsp 1437
 Db 4831 TGTCAAAAGCTTAAGAACTAGCTCTAGAGGCTCTACTGAAAGCAAGAAAGTTAGTGAA 4890
 Qy 1438 GluIleGluValLeuLysAlaGlnLeuLysAlaIleGlu-----GluArg 1452
 Db 4891 GAAATTAATCTTTGAAATCTTCT--AAAGTTGCAGAAAGTACTGACTGCGCAAGAGAA 4947
 Qy 1453 LeuGluIleLysAspArgAspTyrPheGluLeuValGlnThrAlaAsnThr----- 1469
 Db 4948 CACAAGAGAGCTACAAAGAGAGTATGAATCTCTGCACTCCTTATGCAATGTTAGTAAT 5007
 Qy 1470 -----AsnLeuValGlu----- 1473
 Db 5008 GAAGCAGAAAGATTCAGCATGTGTGTGAAGCTGTGAGGCAAGAGAAACAAAGACTGTAT 5067
 Qy 1474 GlyLysLeu-----GluThrProLeuGlnAlaAspHisGlu 1485
 Db 5068 GGCAGAGTTAAGAACACAGAGAGGCAACACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5127
 Qy 1486 GluAspSerIleAspArgArgSerGluMet----- 1496
 Db 5128 GAGCAGAAAGAAATGAGAGAAATGAAGAAAGATGTGTAATCTTAATCAACAGAGAG 5187
 Qy 1497 -----GluIleLysValLeuGly 1502

Db 5188 AAAATCTAGAGCTGGAAGAGAGATGACCGGCTTAGGCGAGAGAGTCCACCTGACAGA 5247
 Qy 1503 GluLys-----LeuGluArgAsnGlnTyrLeuLeuGluArg 1514
 Db 5248 GATACAGCTAAAGAGCTGATGGAACACTCTTTCTTCCAAATGCCAGCATAGAGAGAA 5307
 Qy 1515 LeuGlnGluGluLysLeuGlnLeuSerAsnLysLeuGluIleLeuGlnLysGluMetGlu 1534
 Db 5308 CTGAAAGGCTCAAAATGAG-----TATGAACCCCTTTCTTAAGAACTTTAGAG 5355
 Qy 1535 ThrSerValLeuLeuLysAspAspLeuGlnGlnLysLeuGluSerLeu-----Leu 1551
 Db 5356 TCTTAATGTGTGAAGAAAGCTCTTAAGTGAAGAGCTTCAAGATTTAAAGCATCAGATA 5415
 Qy 1552 SerGluAsnIleIleLeuLysGluAsnIleAspThrThrLeuLysHisAspArgThr 1571
 Db 5416 GAGATTAATGTATCTTAACAAAGCTTAACCTAGAGGCCAGAGAAACATGATACCAACG 5475
 Qy 1572 GlnAlaGlnLeuGlnLysThrGlnGln-----GluLeuGlnLeuAlaLysAsnLeu 1588
 Db 5476 AATGCACATGAAGAGGAAACACAGCTATACAGGTGAGAGCTGAAGACAGACAGCTCTG 5535
 Qy 1589 AlaIleAlaIleSerAspAsnCysProlIleThrGlnGluLysGluThrSerAlaAspCys 1608
 Db 5536 AGTATGACACAGAGACCTACATGT-----TCAGAAATGCGTTCCATCAGCCAAAGAGT 5586
 Qy 1609 ValHisProLeuGluGluLysIleLeuLeuThrGluGluLeuHisGlnLysThrAsn 1628
 Db 5587 GCCAACCTCGCTGTAAAGATTTCACTCAGCTCATGTGTAATTAAT----- 5634
 Qy 1629 GluGlnGluLysLeuLeuHisGluLysAsnGluGlnAlaGlnValGluLeuLys 1648
 Db 5635 -----AACTACCTACAGACAGATTCAT---CAGCTCAAA 5664
 Qy 1649 CysGluValGlnHisLysMetLysSerMetIleGluSerLysSerSerLeuGluSerLeu 1668
 Db 5665 GAAAGAAATGCTGATGTAG 5724
 Qy 1669 GlnHisGluLysHisAspThrGluGlnGlnLeuLeuAlaLeuLysGlnGlnMetGlnVal 1688
 Db 5725 GAAATGAGAAATAATCTTCTACAGTACAGATTCACAAAGATGTGAACTTAATAATG 5784
 Qy 1689 ValThrGlnGluLysLeuLeu-----GlnGlnThrHisGlnHisLeuThr 1704
 Db 5785 CTTCAGAGAGAGATGAACCAAAATGAACCTGTTAATTCAGCAAAATCCAAAGAAAGTCC-- 5841
 Qy 1705 AlaGluValAspHisLysLeuGlnAsnIleGluLeuGluLysAsnPhelLysAsnGlnAla 1724
 Db 5842 TCCAGAGCTTCAAACTAAGAGAGACAGCAGAA----- 5874
 Qy 1725 GlnGlnLysThrThrLysGluGlnCysLeuLeuAsnGlnValGluLeuGlnGlnSer 1744
 Db 5875 GAGAGAGAGAGATGATTGGAAGAGGCTTATGAAATCAATTAAGCAAGAAATGAAGAGC 5934
 Qy 1745 GlnHisArgLeuGlnCysGluIleGluLeuMetLysSerLeuLysAspLysGluSer 1764
 Db 5935 -----ATTGCGAAATTAAGTCTCAGAGATGTATACAGATGCCCAATA 5973
 Qy 1765 AlaLeuGluThrLeuLysGluSerGluGlnLysValIleAsnLeuAsnGlnGluMetGlu 1784
 Db 5974 AAAATAGAGCTATGTG--GAATCTGAAATGAAG-----AACCTTAAGAG----- 6015
 Qy 1785 MetValMetLeuGluMetGluLeuLysAsnSerGlnArgThrValIleAlaGluArg 1804
 Db 6016 -----TGTGTAGTGAATTTGAAAGAGAAAGAGAGAGAGAGAGAGAGAGAGAG 6063
 Qy 1805 AspGlnLeuGlnAspAspLeuArgGluSerValGluMetSerIleGlu----- 1820
 Db 6064 ACTAAGCTGGAATCAGAAATTCAGAAAGAAATATTTGGAAGAAATACAAAGAGCTCAGAAA 6123
 Qy 1821 -----ThrGlnAspAspLeuArgGluAlaGlnGlnAlaLeuGlnGlnLysAsp 1837
 Db 6124 GAACCCGAAATTAAGAGCCATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6183

QY 1838 LysValGlnGluLeuThrSerGlnIleSerValLeuGlnGluLysIleSerLeuGlu 1857
 Db 6184 GAAGTAAAGAGCTACAGAGAGAGCTGATCAGATATCAAGAGAAATAGGCTCGAG 6243
 QY 1858 AsnGlnMetLeuTyrAsnValAlaThrValLys-----GlnThrLeuSer 1872
 Db 6244 AGA-----ACTGTTAAAGCTTAGAATTTGTTCAAACTGATCT 6282
 QY 1873 GluTrpAspSerLeuAsnGlnSerLysGlnIleLeuPheSerGlnIleGlu 1889
 Db 6283 CAAAAA---GATTGGAAATTAACCAAGAAAAAATCTGGCTCAAGAGTTGAACACCGCA 6339
 QY 1889 ----- 1889
 Db 6340 AAGCACAAGCAGATTAGCTAGCTCAAGCTCTGCTAGATACACTCAAGTGAAGCA 6399
 QY 1890 -----ThrLeuSerLeuSerLeuLysGlnLysGlnPheAlaLeuGlnAlaGluLys 1907
 Db 6400 GCAGAGGCTCTAGCAGACATCTCAAGTTGAAAAAGAGACTGATCAATAATTAAGATCA 6459
 QY 1908 AspLysAlaAspAlaAlaTrpLysThrIleAspIleThrGlnLysIleSerAsnIleGlu 1927
 Db 6460 GTTAAAGCCAGATGAAGAAAGAGATGAAGATCTTGAGCGAGAGAGTGAACAGCGCAGA 6519
 QY 1928 GlnGlnLeuGlnGlnGlnAlaThrAsnLeuLysGlnThrLeuTyrGlnLysGlnSerLeu 1947
 Db 6520 GAGAGACACTGAAAGAGAGAGATATGCAAGAGAACTG-----GATGCTTG 6570
 QY 1948 IleGlnCysLysGlnGlnIleAlaLeuAsnThrGlnIleAsnThrGlnIleLysSer 1967
 Db 6571 CGCAGAGAAAAAGCCACTTGGAAGAGACAAATGAGAGATTCAGCTTCACTTGAACAG 6630
 QY 1968 LysAspLeuAlaLeuGlnLysMetGlnGlnGlnLysAspGlnAlaAlaAsnLysValIle 1987
 Db 6631 AAAGACAGAGAACTCAGCAACTTCAGGAAAACTTGACAGCTGCTGACCCAGCTTGA 6690
 QY 1988 AlaLeuThrGlnLysMetSerSerLeuGlnGlnIleAsnGlnLysValIleThrLeu 2007
 Db 6691 GCCTTACTAGAGCATGCTCTCCCTCCAGATGATCGTGACAGGATGATGATGAGCT 6750
 QY 2008 LysGlnGlnLys-----GlnLysGlnLysThrPheTyrLeu 2019
 Db 6751 AAGAAATGGAGAGAGAGTTAGTGTGATGATCAAAAGAGAGAAATTAAGATC 6810
 QY 2020 Gln-----ArgProSerLysGlnGlnAsnSerSerSerGlnMetGlu 2032
 Db 6811 AAGAGATTAATTCAGATGTTCTTAAGATCAACTTACACATGATCCATCATATGAA 6870
 QY 2033 GlnLeuArgGlnSerLeu----- 2038
 Db 6871 GAATTAAGATTAATTCAGATGTTCTTAAGATCAACTTACACATGATCCATCATATGAA 6930
 QY 2039 -----LysThrLysAsp 2042
 Db 6931 CAGACAGAGGTCACGCTCAGCAGAGGCTGTGATCTCTACAGGGGAAAAACAAGAA 6990
 QY 2043 Leu-----GlnLeuGlnGlnAlaGlnLysGlnIleSerGlnAlaThrAsnGlnLys 2060
 Db 6991 CTTTGTCCACGCTAG 7050
 QY 2061 AsnLeuThrAlaLysIleSerSerLeuGlnGlnLysIle-----LeuGlnAsnAla 2077
 Db 7051 AAGTGGATACAGAACTTAAGAGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7110
 QY 2078 -----SerIleLeuAsnGlnAlaValSerGln 2086
 Db 7111 GAAAAATGTAGCAAAAAAGAACTTGAAGAGAGATCAATGAGGAGCAAGAGAGAGAG 7170
 QY 2087 ArgGlnAsnLeuArgHisSerLysGlnGlnIleValSerGlnLeuGlnLysLeuSerLeu 2106
 Db 7171 ATTCAAAATCTAGTTCAGTTATGAACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7227

QY 2107 ThrLeuLysSerArgAspHisAlaPheAlaGlnSerLysArgLys----- 2121
 Db 7228 GAAGTACACAGTAGAGCTGATGAGAAATTAATGAAAGAGCAAAAAAGATTAAGCTCG 7287
 QY 2122 -----LysAspGlnAlaValAsn-----LysIleAlaSerLeuAlaGlu 2134
 Db 7288 CTTTGTCCAG 7347
 QY 2135 -----GlnLys-----IleLeuThrLysGlnMetLysGln-----Phe 2146
 Db 7348 AAAGAAATTAAG 7407
 QY 2147 ArgAspSerLysGlnSerLeuGlnGlnIleSerSerHisLeuSerGlnLysCysThr 2166
 Db 7408 GAAAG 7467
 QY 2167 TyrLysThrGlnLeuGlnMetLeuLysGlnGlnLysGlnAspIleAsnAsnLysLeuAla 2186
 Db 7468 ATCAAA-----AAGGAAACATTCAGCAAAAG----- 7494
 QY 2187 GlnLysValLysGlnValAspLysLeuGlnGlnIleSerSerLeuLysGlnGlnLeu 2206
 Db 7495 -----GCAAGATGAGATCTTGTGTAATCCATGCTCTCTCCAAATGATGCA 7545
 QY 2207 AspGlnIleGlnMetGlnLeuArgAsnGlnLysLeuArgAsnTyrGlnLysCysGlnLys 2226
 Db 7546 GACCGATAGTGGGTGATCATCAACAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7596
 QY 2227 MetAspIleMetGlnLysGln-----IleSerValLeuArgLysMetGlnAsnGlnPro 2244
 Db 7597 -----ATCTGGAAAAAGCAACACTCAACAGAGCTGCTGAGAGAGAGAGAGAGAG 2264
 QY 2245 GlnGlnGlnGlnLysAspValAlaGlnArgMetAspIleLeuGlnSerArgAsnGlnLys 2270
 Db 7651 AAGAGAGAAATACAGAGGCTGAGAGAGATGATGATGATGATGATGATGATGATGATG 7710
 QY 2265 Ile-----GlnGlnLeuMetGlnLysIleSerAlaValLysSer 2277
 Db 7711 CTAGATGAG 7770
 QY 2278 GlnGlnIleThrLeuSerSerSerSerSerSerSerSerSerSerSerSerSerSer 2297
 Db 7771 AGCCAAACAAAG 7827
 QY 2298 LysHisCysMetLeuAsnLys-----GlnSerLeuSerSerThrLeuSer 2313
 Db 7828 AAATATCTTAATTAAG 7887
 QY 2314 ArgSerPheGlnSerLeuGlnThrGlnHisValLysLeuAsnThrGlnLeuGlnThrLeu 2333
 Db 7888 AGGTCCCTTAATGAG 7947
 QY 2334 -----LeuAsnLysPheLysValIleTyrArgThrAlaAlaValLysGlnAsp 2349
 Db 7948 AAATATCTTAATTAAG 8007
 QY 2350 -----HisSerLeuIleLysAspTyrGlnLysAspLeu----- 2360
 Db 8008 GAGCTATATCAATCCACCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8067
 QY 2361 -----AlaAlaGlnGlnLysArgHisAspGlnLeuArgLeuGlnGlnLysLeuGln 8127
 Db 8068 TTTTCTCTCTCAAAAG 8184
 QY 2380 HisGlyArgLysTrpSerSerSerSerSerSerSerSerSerSerSerSerSerSer 8199
 Db 8128 GAAGCTCCCAAG 8244
 QY 2400 LeuAsnGlnLeuLeuPheLysLysAlaAsnIleLeuGlnValGlnAspPheSer 2419
 Db 8185 CTTATCATGATGATGAG 8244
 QY 2420 GlnValGlnValPheLeuAsnGlnValGlySerThrLeuGlnGlnLysHisLys 2439

```

Db 8645 GAGGACAGAGAGATTTGGTGGATGGAGAAAGAAATTAACATGCTGACCAAGAAAT 8304
OY 2440 LyselypheheltIntPleugluphegly-----AspleuhisValasp 2455.
Db 8305 AAAGGTCTCAGACAAACAACTCTTTGGAGGTCTATGAGTCTCTCAAAATAGT 8364
OY 2456 AlalysleusergluglymeclnglunlsnhrargyleaSerThrIlegl 2475
Db 8365 AGAGATATGCGCATGAGGAACCTGATGAATAAGAAATATGATGCGAGCTGAG 8424
OY 2476 LeuleuthrlysrargyleaValaValaGlnSerlyIleGlnargyleVal 2495
Db 8425 GAATTGGCA-----CAGTGAAGAA-----CAGGACTCTTAAACAGAGAGAGATGCT 8475
OY 2496 Tytleuansnglnpneglualalysleuglunlsylsuglunlsnlygluleumet 2515
Db 8476 CTCTCTTCTGAAACCGCTTTTCATGAACTCCACTGAGGAAATAGCTTCTCACCTT 8535
OY 2516 ArgargmetgluHshlslyProseralaserValmetgluglunlsnAlaArgleu 2535
Db 8536 GAGAACTTAAACCAAG-----CTCTATCCAAAGATGAGCAATTG 8577
OY 2536 LeuglylleuylsrthrValaGlnaspIuserlyslsleuglnserargyleus--- 2554
Db 8578 CTTCACCTTGCTCCACAACTAGAGATTTCTATAAACAGTGCAGTCTTTCCAGGCT 8637
OY 2555 -----metleuglunlsn-----leuValylsaspAla 2567
Db 8638 ATGGCCAGTCTCCAGAACTGAGAGATCACCTGGAATGAGTGGAGAAATTCGAAAG 8697
OY 2568 MethlsylsuglylulysValaIleleuglnaspIuserArgAsnAlaglu 2587
Db 8698 TCAGAGAGAGGAGAGAGAGGTCTGCACTTCC-----ACAGGCCA 8745
OY 2588 AlaGluleuansnAlametGlnValylsleuthrlyslsGlnaspAsnleuGlnAla 2607
Db 8746 GCTGAAGTACAGAGATTTAAAAAAGCTATGCTTCCACTCCAAATGACAGAGACTA 8805
OY 2608 MetylsugluleuansnleuGlnlymetValalysglValaValProtyllysglu 2627
Db 8806 CTGAAGAAATTTGAAGATCTGCACAGCAATAC-----TTACAGATTAATCAA 8853
OY 2628 GluIleaspAsnleu-----LysThrlys 2635
Db 8854 GAGATCACTGATACATCACTCAAGGCTCACTCAGAGTATCAAGATTAACAAAA 8913
OY 2636 ValValylsIleglumetGluylsIlelysrlySerlysalathrAspGln---GluIle 2654
Db 8914 GCATTCAGATTATGCAAGAGAGAGCTCAGCAGAAAACTCTCTGCGACATGAGCTG 8973
OY 2655 AlaTytleuys----- 2658
Db 8974 CATTCACTCAGAGATGAGAAAGATTCCTGGGAATATCATGAGAGAGAAAGCAAGCAG 9033
OY 2659 -----SerCysleuglnuaspIysgluglyleuargyleuysgluIleuarg 2676
Db 9034 TACCTTATGGCTATCTCAGATTAAGATCAGAGAGCTGATCTGCAGAACTTATAAG 9093
OY 2677 ArgAlaGlnAlaaspAsnAspThrThrValCysValProlysaAspTyrglnlysalSer 2696
Db 9094 GAATTGGAGTCTCTCTCCACAGACTCAGCTCTCAAGAGCAATACCAAGACAGGCA 9153
OY 2697 ThrpheProValThrlysgly-----glylyserglyIleValGlnserThra 2713
Db 9154 TCC-----CCAGAGACATAGCTTCCCAAGATGGTCAAAAATCTGGTTATGAGACGAA 9210
OY 2714 MetleuValleuglnserGluylsAlaAlaAlaGlnleuargIleuSerHslylyls 2733
Db 9211 CTCTCAGAGACCAAGCTCATGACAGCTTA-----AAGGAATTCACCAAAAGAGCTTA 9264
OY 2734 LysTyrlHshlsleuserArgthrMetSerSerSerlyuaspArglylsThr----- 2751

```

```

Db 9265 AGAATTCAACAACGTGACAGCAACTTCTCAGTACTGAAAGAGAAAAACACCTTCC 9324
OY 2752 ---LysAlalysSerAspAlaHisSerHsThrlySerSerHsArglySer--- 2769
Db 9325 ATTCAGCTCTCGATVACAGCTCAGTCTGAGAAACAGACAGCACTATGCTGACCTT 9384
OY 2770 -----ProHslysrThrGlnThrlyThrArgHsGlyProValThr 2782
Db 9385 TTAATCACTGTGCACTTGTGGAGAGAGCTTCAAGAGCTGACAGCGCGGCGCAATAT 9444
OY 2783 ProGluArgSerGlu---metProserLeuHsleuglySerProlysaSerGluSer 2801
Db 9445 ATAGATGTGCTCCAGAGAGCTCCCAAGAAAGATGAGTTCACAGAAAGGTGACCTT 9504
OY 2802 SerThrlysrArgValaValaSerProAsnArgSerGluIleTyrglnleuValmetSer 2821
Db 9505 GAGGAACTAAGGAAACCCAGACAAAGCTTTCTGAAAGTCAAGACAGCACTATGCAACAC 9564
OY 2822 ProGlylysrThrGlyMethHslysrHsIleleuserProserlysr----- 2836
Db 9565 AGACAGAGAGTGAATTAAGAGAGCTGCGAAGAAAGACAGACCAAGAGTGGCT 9624
OY 2837 -----ValGlyleuHslysrValaAlaAlaLeuSerProAsnArgSerGluMet 2852
Db 9625 GCTGAGAAATGCTCTCTGTGCGCCAGAGAGATCAGAGCTTATGACAGAGTGAATG 9684
OY 2853 ProThrGlnHsValIleSerProGlylysrThrGlyleuHslysrAsnleuThrGluSer 2872
Db 9685 GACCTCTCCGAGACTCTCATATGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 9735
OY 2873 ThrleuPheaspAsnleuserSerProCyslysrGlnlyls 2886
Db 9736 CTGTTAATGATCTTACAAAGCAACAGCTTGTGAAAGACCCG 9777

```

RESULT 4

```

US-09-960-253-163
; Sequence 163, Application US/09960253
; Patent NO. US20020123619A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Monamath, Raedoh
; APPLICANT: Lodes, Michael J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.556
; CURRENT FILING DATE: 2001-09-20
; CURRENT APPLICATION NUMBER: US/09/960.253
; NUMBER OF SEQ ID NOS: 187
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 163
; LENGTH: 10096
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-960-253-163

```

Alignment Scores:

```

Pred. No.: 1,15e-64 Length: 10096
Score: 1152.00 Matches: 740
Percent Similarity: 37.41% Conservative: 563
Best Local Similarity: 21.25% Mismatches: 1160
Query Match: 7.80% Indels: 1022
DB: 10 Gaps: 140

```

US-09-150-867-1 (1-2954) x US-09-960-253-163 (1-10096)

```

OY 81 PhealatyrglylnThrSerSerGlylysrThrlyThrMetmetGlyThrProAsnSer 100
Db 465 TTCAGAGAGGACAACTGATTCAGGCNA----- 494
OY 101 LeuglylleleProGlnAlaIleGlnluValPheylsIleIleGlnluIleProAsn 120
Db 495 -----AAACAAATAGAAAAAAGTGAAGCAAGAACTTAAAG 530

```

QY 121 ArgGluPheLeuLeuArgValSerTyrMetGluLeuTyrAsnGluThrValLysAspLeu 140
 Db 531 TGTAAATTCAGCTTGAAGAAAGCCAAAGCTGCGAGCTGCAGATGCTCTCTGAT 590
 QY 141 LeuGlyAspAspArgArgLys-----LysProLeuGluLeuArgGluAspPheAsn 157
 Db 591 CCATGCATATCACCAACAAAAATTTTACACTCCACTACCAACCAAGCAATATATAGT 650
 QY 158 ArgAsnValTyrValAlaAspLeuThrGluGluLeuValMetValProGluHisValIle 177
 Db 651 GGTTCCAAAGAT---GAAGATCTAAAGAAAAATATATATAAAGAGTTGAA----- 698
 QY 178 GluThrPheLysGlyGlyLysAsnArgHisTyrGlyGluThrLysMetAsnAspHis 197
 Db 699 -----GAACGAAAAAGATTAGAGCAGCAGAGGTAAACCTTCAGGCT 740
 QY 198 SerSerArgSerHisThrIlePheArgMetIleValGluSerArgAsp-----ArgAsn 215
 Db 741 AAAAAGCAACCCAGACTCTCCACAGCCACATGATACCCAGCAGATTCGCCGCGCAT 800
 QY 216 AspProThrAsnSerGlu---AsnCyAspGlyAlaValMetValSerHisLeuAsnLeu 234
 Db 801 CAGGCTTCATCTGCTGCTCTCATGCGACAGCAGAGAAAGCCCAAGTCATCTTCA--- 857
 QY 235 ValAspLeuAlaGlySerGluArgAlaSerGluThrGlyAlaGluGlyValArgLeuLys 254
 Db 858 -----TCTAATVTCCAAGAACT----- 875
 QY 255 GluGlyCysAsnIleAsnArgSerLeuPheIleLeuGluValIleLysLysLeuSer 274
 Db 876 -----CCAAATTAGAGAGATTCCTCT 896
 QY 275 AspGlyGluAlaGlyLysPheIleAsnTyrArgAspSerLysLeuThr----- 290
 Db 897 GCATCTTACTTTCTGGGAACTAGAGTGACTCCAACTGCATCACTTCCTCAAAATAGC 956
 QY 291 ---ArgIleLeuGluAsnSerLeuGlyGlyAsnAlaLysThrValIleLysThrIle 309
 Db 957 AAAAGAGATCTAATAGCAGTTCTTGGCAATTCTAC----- 995
 QY 310 ThrProValSerPheAspGluThrLeuSerThrIleGluPheAlaSerThrAlaLysHis 329
 Db 996 AGCCCTCATCTTTGAT-----CAATTAAAGCCGACAGATCAAG 1037
 QY 330 ValArgAsnThrProHisValAsnGluValLeuAspAspGluAlaLeuLeuLysArgTyr 349
 Db 1038 CTAAGAAAC-----AAGATTATGAGTTG-----GAACCTGCGCTGCAAGACAT 1082
 QY 350 ArgGlyGluIle-----LeuAspLeuLysLysGluLeuLys--- 361
 Db 1083 GAAAAAGAAATGAAGGCCAAAGTAATTAAGTTCAAGAACTCCAACTCCAACTGGAGAA 1142
 QY 362 ---AsnLeuGluSerSerSerGluThrLysAlaGluAlaMetAlaLysGluGluHisThr 380
 Db 1143 GCAAAAGTGGAAATTAATGAAGAAAGAGTTTGAACAAAGTGATGAGATCAAGAG 1202
 QY 381 GluLeuLeuAlaGluIleLysGluLeu-----HisLysGluArgGluAspArgLys 397
 Db 1203 AGAACACACGACATAGCAGCCGCTCAACCAAGTATCTCATTTGCAACAAAAGCTG 1262
 QY 398 ThrHisLeuThrAsnIleValAlaLysSerGluGluLysGluGln----- 413
 Db 1263 AAAAATTTGACGAGATAGTTGAGTTCAGCGACAAATATGACAAATGCGCAATGTCT 1322
 QY 414 ---AspGluArgValLysArgLysArgValThrThrAlaPro---GlyLysIleGln 431
 Db 1323 CTGGAACGAAATTAAGAAAAAGAAAGAGTTTCAAGAGAGCTCTCCGCAACAG 1382
 QY 432 AsnSerLeuHisAlaSerGlyValSerAspPheAspMetLeuSerArgLeu---ProGlyLys 451
 Db 1383 CGTCTCTTCCAAACACTGCAGCAGAGTGATCCAGATGAAGGCAAGCTACCCAGAG 1442
 QY 451 nHeserLysLysAlaLysPheSerAspMetProSerPheProGluIleAspAspSerVal 471

Db 1443 TTAACAGCAG-----CCAAAGAT 1460
 QY 471 LysThrGluPheSerAspPheAspAlaLeuSerMetLeuAspSerAsnGlyIleAs 491
 Db 1461 ATGCACACAGCTCTGAGGCTGAGATGATTAACATCAATCAATGAAACCAACAGCTGAA 1520
 QY 491 PAlaGluThrAsnLeuAlaSerLysValThrHisArgGluLysThrSerLeuHisGlnSe 511
 Db 1521 AACAAATTTGAGACGTTTAAAGCAAAAGTTGTCG-AGAGTGAACAGCGCTTCCAGCGAG 1579
 QY 511 rMetIleAspPheGlyGluIleSerAspSerValGluPheHisAspSerSerLysGluAs 531
 Db 1580 TCAGATCAAGAGAAATGAGCTGAGAGAAAGCATGAG-----GAATGACAGAGAAA 1633
 QY 531 nGluLeuGluThrLeuProLysAspSerGlyAspMetAlaGluCysArgLysAlaSerPh 551
 Db 1634 CAACCTCTTAAGATGATCACTGAGCAAAAGCCAGAAAGTCTGCCACTGAGAGCGA 1693
 QY 551 eGluLysGluIleThrSerLeuGluGluGluGluGluGluGluGluGluGluGluGlu 571
 Db 1694 ACTCAAGACATCAAAACAGTGTAAATCAAGACCCAGAAATTTGCAAGAAATGAAAC 1753
 QY 571 uLeuValGluSerPheGluLeuLysIleAlaGluLeuGluGluGluGluGluGluGlu 591
 Db 1754 GAAGATACCTCTGAGGAAACCATGTTAAGAGATCTTCAAGAAATA----- 1802
 QY 591 aLysAsnLeuGluMetValThrAsnSerArgGluHisSerIle----- 605
 Db 1803 -----AATCAGCAAGAAACCTCTTGCCTTGAACAAACTGAA 1840
 QY 606 -----AsnAlaGluValGluThrAspValGluLysGluValAlaArgLysGlu 621
 Db 1841 GCTTCTGCTGCTGATCTGGAAGAAACAGCAGATGTTCTCAAGACCTTTGAAAGAAAG 1900
 QY 621 uMetSerValLeuGluLysAspSerGlyTyrAsnAlaSerAsnSerPheGluAspSerSe 641
 Db 1901 A-----GAACATCACTTGAACCACTTAATATATATAGTTAAGCAAGACAGA 1945
 QY 641 rValAspGlyLysArgLeuSerSerSerHisAspLysIleGluHisArgLysMetLe 661
 Db 1946 GAAAGGTCCAAAGCCTTCTGAGTCT-----TT 1975
 QY 661 uGluGluLysIleValAspLeuGluGluPheIleGluAsnLeuAsnLys----- 677
 Db 1976 AGAGTAAAGAAAGAAATATCAAGAAATGAAGAAAGAAAGAAACTGTTTCTTCTG 2035
 QY 678 -LysSerGluAsnAspLys-----GluLysSerSerGluGluAspPheMetGluSe 694
 Db 2036 GAAAGTGAAGAAAGAAAGCTTTTAATCACTGAGATGAAATCAAGAAAGAAAGCTTCAAG 2095
 QY 694 r-----IleGluLeuGly----- 698
 Db 2096 TAAATTAATCACTGTAAGCTTGTGAAGACACAGCAAAATTAAGTCAATGAATACAA 2155
 QY 699 -----GluAlaIleMetAlaGluLysAlaAsnAlaLeuGluGluLeuAlaLeuWe 715
 Db 2156 CGAGAGATGAAGAGCTGAGATGAGACAGAGAAACCTTAAGTGCAGATCAAGAACT 2215
 QY 715 tArgAspAsnPheAspAsnIleIleLeuGluAsnGluThrLeuLysArgGluIleAlaAs 735
 Db 2216 TCACAACTGTTAACAAGTAATGATGAGAGTAGAGACCCAGAAACTAGCTTATATGGA 2275
 QY 735 pLeuGluArgSer----- 739
 Db 2276 GCTTACAGCAAGAAAGCTGATCTCAATCAAGAAACATCAAGAAATAGAAATATGTTG 2335
 QY 740 -LeuLysGluAsnGluGluThrAsnGluPheGluIleLeuGluLysGluThrGluLysGlu 759
 Db 2336 TTGGAAGACTTCTGAGCTTACTGGCAAGTTGAAGATCTA-----GA 2377
 QY 759 uHisGluAlaGluLeuIleHisGluIleGlySerLeuLysLysLeuValGluAsnAlaGlu 779

Db 2378 ACACAGCTTCAGTTACTGTCA-----AATGAATTAATGACAAAGACCG 2422
 QY 779 UMETTYRAsnglnAsnleuGlunAspLeuGluthrLysThrLysleuLysgluGl 799
 Db 2423 GTGTAC---CAAGACTTGATGCCGAATATAGAGCCCTCAGGATCTGTAAATCCAA 2479
 QY 799 nglnlLglLnu----- 803
 Db 2480 AGATGCTTCTCTGTGTGACAAATGAAGATCATCAGAGAACTTTGGCTTTGATCAGCA 2539
 QY 803 ----- 803
 Db 2540 GCCTGCATGCATCATCTCTTGCATAATATATATGAGAACAGAGCATCCCTCAGA 2599
 QY 804 ----AlaGlulEuArgLysArgAlaAspAsn-----LeuGlulY 815
 Db 2600 GAGAGAGGAATGTGCTGTAGAGAGACCAAAAGTCGAAAAATTCGCCATCTCACAAA 2659
 QY 815 sLysValArgAsnPhaAsnPhaLeuSerValSerMet-----GlyAspSerGI 830
 Db 2660 TAGAGTGTATTCACCTTGAATTTTTCATTAGAGTCTCAAAAACAGATGAAGCTCAGACCTGCA 2719
 QY 830 uLysLeuCySgluGlulPheGlnLeuLysGlnSerLeuSerAspAlaGluAlaValTh 850
 Db 2720 AAACGACTGTGAAGAGTTGGTGAATCAAAAGAGAAATAGAA-----GAAATTCAT 2773
 QY 850 rArgAspAlaGlnLysgluLysSerPheLeuArgSerGluAsnleuGlulEuLysgluL 870
 Db 2774 GAAAGCAGAACAGATGCATCAAAAGTTTGTGCTGAAACAGTCAGCGCATTAAGATT 2833
 QY 870 smetGluAspThrSerAsnTrpTyrAsn-----Glu 880
 Db 2834 ACAGAGAACACACTCTCTGCTACCAAGATGTGTGCTGAAACCTTAAGTCCCTTGAGAA 2893
 QY 880 nLysGluLysAlaAlaSerLeuPheGluLysGlnleuGlulThrLysSerAsnTyrLy 900
 Db 2894 CAAGGAAAGAGCTGCACACTTTTAATGATAGTAAGAACTGACGAGCAGAGATGCA 2953
 QY 900 sLysMetGlu-----AlaAspLeuGlnLysgluLnu----- 911
 Db 2954 AGAATTAAAAAGACCAACCATCTTACTTGAGACTCTCTTAAGAGCTACAACTTTATTC 3013
 QY 912 -----SerAl 913
 Db 3014 CGAAACCTTAAGCTTGAGAGAAAGAAATGATTCATCATTTCTTAATATAAGGGA 3073
 QY 913 apheAsnGlnlLeuAsnTyrLeuAsnGlyLeuLeu-----AlaGlyLy 927
 Db 3074 AATTGAAGACTGACCCAGAGATGGACTCTTAAGAAATTAATGATCTCTTAATCA 3133
 QY 927 sValProArgAspLeuLeuSerArgVal-----GluLeuGI 939
 Db 3134 AGACAGATGAACTTAATCCAGAAAGTGAAGATTGCAAACTATATATAGTGAAGGGA 3193
 QY 939 uLysLysValSerGluPheSerLysGlnleuGlu----- 950
 Db 3194 GAAAGACATTTTCAGAGTTATCTGATCAGTACAGCAAGAAAATTTTACTACAAAG 3253
 QY 951 -----LysAlaLeuGlul 955
 Db 3254 ATGTGAAGAAACCGGAATGCATATGAGATCTTATGTCAAAATATACAAAGCAGCAGCA 3313
 QY 955 uLysAsnAla-----LeuGluAsnGlnValThrCysLeuSerGluTyrLysPh 971
 Db 3314 AAAGAAATTTCTTAATTAAGATGCTTGTCTTAATGAATGCACTGCTTTGTGAAAAATAG- 3371
 QY 971 eLeuProAsnGlnValGluCysLeuLysAsnGlnlLeuSerLysAlaSerGluGluLeu 991
 Db 3372 -----AAAAATGAGTTGGAACAGCTAAAGGAAGCATTTGCAAGAGAACACCAAGATCTT 3427
 QY 991 tLeu---LeuLysGlnLugluLysSerAlaSerlleLysSerLysGlnlulLeI 1010
 Db 3428 AACAAATTTAGCATTTTCTGAGAGAAAGAAATCAGAAATCTGATCTAGAGTTGAGAGAGT 3487

QY 1010 eMetGln-----GluGlnSerGlu----- 1016
 Db 3488 GCAGAACCTCTGAGATCTGAGATGACAGATACCAAAAAATTTCAAGACGAGCTGCG 3547
 QY 1017 -----GlnlleuGlnleuThrAspGluValThrLysThrGlnSerLysVa 1032
 Db 3548 TGGTTAAAGCAAAATCATGACTTTAAAGAGAACAAAAACAAATATGCAAAAGAGAGT 3607
 QY 1032 1-----GlnGlnThrGluGlnLugluLysMetLysAsp-- 1047
 Db 3608 TAATGACTTATTAACAAGAGATGAACACCTGATGAAGATTAATGAAGACTAAACATGAATG 3667
 QY 1048 ---AspLeuPheGlnLysTyrlleArgAsnLysSerGluAlaGluAspLeuLysgl 1066
 Db 3668 TCAAAATCTTAATTCAGAACCAATTTAGAACTGTGTGAAGAAAGAGAGTGAAGAA 3727
 QY 1066 uMetGlnAsnleuLysgluThrMetGlu----- 1075
 Db 3728 TCATGTATTTTAAACCTCAGATGATCTTGAAGCTTAAGAAATTTCTCATAGTTA 3787
 QY 1076 -----SerValGluValLyslleAlaAspThrLyslleGluLuglu 1090
 Db 3788 TAATGCCAGTTGTGTCATTAAGAGTATGCTTAAGAAATTAAGAAATTAACCTTACAGA 3847
 QY 1090 uThrlleArgAspLysgluGlnleuLyslleGluLysLysTyrPhePheGlnAlaMetGI 1110
 Db 3848 AAGTGAGAGAGAGAGAGTCCCTGCACATGA-----TTACA 3886
 QY 1110 nThrillePheProIleThrProLeuSerAspSerLeuProSerLysleuValGluGI 1130
 Db 3887 GACATTT-----AGAGAGATCTTAACCAACCAATTTGCAAGACAT 3928
 QY 1130 yAsnSerGlnAspProIleGluLleAsnAspTyr----- 1141
 Db 3929 GCAGTCACAAAGAAATTAAGTGCCTTAAGACTGTAAGATAGATCGGAAAGAAATATAT 3988
 QY 1142 -----HisAsnleuLleAlaLeuAlaThrGluArgAsnlleMetValCysLe 1158
 Db 3989 TTCAGGCCCTCATGAGTTGTCAACAGCTCAAAAGCAGACATGCACCTTCAGTCTCTCT 4048
 QY 1158 uGluThrGluArgAsnSerLeuLysgluGlnVallleAspLeuAsnThrGlnleuGlnSe 1178
 Db 4049 GCAAAACATCAATGAACAAGCTGAATGAG-----CTAGGAAATATATGAAAT 4096
 QY 1178 rLeuGlnAlaAlaSerlleGlu-----LysSerAspLeuGI 1190
 Db 4097 ACTGCAGCTGAAAGATATGAACCTGTAAGTGAATGATTCAGAGCTCAGATGTAT 4156
 QY 1190 nLysProLysGlnAspLeuGlnLugluVallleLysleuLeuGlnleuMetGluLeuLe 1210
 Db 4157 CACAGCAACTAGAAATGGCAGAGAGGTAGGAACTTAAATGAAATTAATATAT 4216
 QY 1210 uLysGly----- 1212
 Db 4217 AAATGATGACAGTGTCTTTCATGTGATTAAGTGAAGACATATACAGAGAGTGAAT 4276
 QY 1212 ----- 1212
 Db 4277 TGGTGAACAAATGAAACAGACCCGTGTCTTTGGCTTCATTTGACGAGAGTAAATTC 4336
 QY 1213 -----HisleuThrAspSerGlnleuSerlleGluLysleuGlnleuGluAsnleuGI 1230
 Db 4337 CTAGAGACACTTGACA-----TTGTCAAGCAAAAGAAATTAATGATGACTTTGCCGA 4387
 QY 1230 uValThrGluLysleuGlnThrleuGlnLugluMetLysAsnlleThrilleGluArgAs 1250
 Db 4388 ATTCAGAGAGAAATTTTATCTTAACAAGTGAACACAAAAATTTTATCATGATCAGCACTG 4447
 QY 1250 nglnleuGlnThrAsnPhaGluAspLeuLysAlaGluLlleAspSerlleLysGlnAspLe 1270
 Db 4448 TCAGATGAGCTCTAAATATGTCAGAGCTGCAGACCTATGTTGATTCATTAAG----- 4499


```

OY 1270 usergluasnileglugluserllegluuthrglnaspgluuauargalialieglucl 1290
    ::::::::::::::::::::
DB 4500 -GCCGAAATTTGGTCTGTGCA-----ACGAATCTGCAAACTTCAAGTGA 4546
OY 1290 uleuargluginlysleuValaspserrheargluginleuLeuaspCysSerVa 1310
    ::::::::::::::::::::
DB 4547 CTTCGTGAGAGAGATGAGCTG-----GGCTTGAGAGAGGGCTGCTCCATCCTGTC 4600
OY 1310 lglYlleserSerProasnHisAspAlaValaAsnGlnGluValSerleuGlycl 1330
    ::::::::::::::::::::
DB 4601 ATCCTTTGTGTGCTTACAGAGCTCTACTTACG-----AGTTTGGAGAGA 4645
OY 1330 uValasSerleuGlnSerGluMetleuArgGlyGluArgAspGluLeuGlnThrSerCy 1350
    ::::::::::::::::::::
DB 4646 C---TCCCTCTTTACAGAGCTCTTTTAAAGA-----CAGACAGAGAGA 4684
OY 1350 slyAlaLeuValSerGluLeuGlu-----1358
    ::::::::::::::::::::
DB 4685 TATGCTCTTTTGGATTAATTTAGAAAGGGCGCTTTCAGCAAAACAGTCAGTGAATGA 4744
OY 1359 -----LeuLeuArgAlaHisValysSerValGluGlyGluAsnleuGlucl 1374
    ::::::::::::::::::::
DB 4745 AGTATTTTCAGACAGCTGCGACACCTATGTTGACTTAAGGCCCAAAATTTGGTCTT 4804
OY 1374 ethrlystysleuasnGlyleuGluGlyGluLeuGlyLysSerGlu-----Gly 1391
    ::::::::::::::::::::
DB 4805 GTCAACGCAATCTGAGAAACTTTCAAGGTGACTTGGTGAAGAGATGCAAGCTGGCTTGA 4864
OY 1391 usergluValleuysSerMetleuGluAsnleuLysGluAspAsnAsnlySleuLys-- 1410
    ::::::::::::::::::::
DB 4865 GGAAGGGGCTGCTTCATCCCTGTCATCCTCTGTGTGCTGACAGCTCTACTTACAGAG 4924
OY 1411 -----GluGlnAlaGluGluThrSer-----1417
    ::::::::::::::::::::
DB 4925 TTTGGAGACTCCTCCTTTTACAGAGCTCTTTTGAACAGCAGAGATATGCTCTTTT 4984
OY 1418 -----SerLyGluasnGlnPheSerleuGlnGluValPhe-----1429
    ::::::::::::::::::::
DB 4985 GAGTAATTTAGAAAGGGTGTGTTTCAAGCAAAACAGTCAGTGAATGAAATTTTGCAG 5044
OY 1430 -----SerGlySerGlnLysLeuVa 1436
    ::::::::::::::::::::
DB 5045 CAGTCTGACGAGAGAAATCTGACCAAGAAACCCCTTGCGCCCAAGCAAGAGGGTGT 5104
OY 1436 laspgluile-----GluValleuysAlaGlnleuLysAlaAlaGluGluAr 1452
    ::::::::::::::::::::
DB 5105 TGAAGCTTGAGCTCCCTCTGAGGTGACCGGAGCTGCCGAGAGAGCTAGAGAGAA 5164
OY 1452 gleuGlu-----lleuysAspArgAspTyrPheGluMetleuValGlnThrAlaAs 1468
    ::::::::::::::::::::
DB 5165 AATGGAAGCTCAAGGATTAATGAAAAATTAAGAAATTCAGAGCTCGAGCACTTATTAG 5224
OY 1468 nThrAsnleuValGlu---GlyLysleuGluThrProleuGlnAlaAspHisGlu--- 1485
    ::::::::::::::::::::
DB 5225 TTCTGAAAGCAAGAGCTTGACTGCTTAGCAAGCAGTATTGTGCAAAAATGAACAGTG 5284
OY 1486 -GluAspSerlleAspArgArgSerGluGluMetGluLysValleuGlyGluLysle 1505
    ::::::::::::::::::::
DB 5285 GCAACAGAAAGCTGACAAAGCTGCTGAGATGAGTCCAAAGTGGCCGCAAAAAGAA 5344
OY 1505 uGluArgAsnGlnTyrleuLeuGlu-----ArgleuGlnGluGluLysleuGlu 1521
    ::::::::::::::::::::
DB 5345 ACAGAGCGAAACACTGTCACTTGAAGCTGGAAGTGAACAGACACTCCAGTCAAGTCTGA 5404
OY 1521 uLeuSerAsnLysleuGluLysleuGlnLysGluMetGluThrSerValleuLysAs 1541
    ::::::::::::::::::::
DB 5405 CTTAAGTTCGCG-----TCTTGTGCTGGCAT 5431
OY 1541 paspleuGlnGlnLysleuGlnSerleuLeuSerGluAsnlelleleuLysGluAsnll 1561
    ::::::::::::::::::::
DB 5432 CGACACAGAAAGATGCTATTCAAGGC-----CGAAATGAGAGCTG 5470
OY 1561 eaSprThrThrleuLysHisHisSerAspThrGlnAlaGlnleuGlnLys-----1577

```

```

DB 5471 TGACATATCAAAAGACATATCTTCAGAAACTACAGAAAGAACCAAGACATGATGTCA 5530
    ::::::::::::::::::::
OY 1578 -----ThrGlnGlnGluLeuGlnLeu-----AlaLys 1586
    ::::::::::::::::::::
DB 5531 TCAGATTGTGATTAAGATGCTGACGAGACCTCAATCTGACATTGAGAAATAACTGA 5590
OY 1586 sasleuAlaIleAlaIleAserAspAsnCySProlleThrGlnGluLysGluThr-----1604
    ::::::::::::::::::::
DB 5591 GACGTGCTCAGTGAACCCACAGAGAGATGCTGGGCAACAGTCCCAATACCAATTA 5650
OY 1605 -----SerAlaAspCysValHisProleu-----1612
    ::::::::::::::::::::
DB 5651 TGAGCTCCAGGGAGATTAACCCAGAGGCTCTTCAGAAATGCAATTTCTGAATGTGAT 5710
OY 1613 -----GluLysLysleuLeu 1618
    ::::::::::::::::::::
DB 5711 TTCTGCTCCTTAATGCTTTGTGATCCTATGATTTCTCTGGGAATCAGAAATATCTATA 5770
OY 1618 uLeuThrGluGluLeuHisGlnLysThrAsnGlnGluLysleuLeuHisGluLysAs 1638
    ::::::::::::::::::::
DB 5771 TCTTCAACTGCGGATAAAGAGACATCAATTAATTTGATGATCTTCTGATGATGA 5830
OY 1638 nGluLeuGlnGlnAlaGlnValGluLeuLysCysGluValGluHisleuMetLysSerMe 1658
    ::::::::::::::::::::
DB 5831 GAACGCTGACAGA-----AAAGTTAAAGTTGCTTAATGAAT 5869
OY 1658 lileGluSerLysSerSerleuGlnSerleuGlnHisGluLysHisAspThrGluGlnl 1678
    ::::::::::::::::::::
DB 5870 GAAAGAAATTAAGCTCAAACTC-----CATTTACAGAGGTGACA 5908
OY 1678 nleuLeuAla-----LeuLysGlnGlnMetGlnValAlaThrGly 1691
    ::::::::::::::::::::
DB 5909 ACTAATGACCAAAATTAAGACATGCAATGAATTTGAAAAAATAGTGGGAACTTAAGAA 5968
OY 1691 nGluLysLysGluLeuGlnGlnThrHisGlnHisleuThrAlaGluValAspHisleuLys 1711
    ::::::::::::::::::::
DB 5969 AGAAACTCAGATTTAAGTGAATAATTTCTGCTGATCACCAGAGACTTCT 6028
OY 1711 sGluAsnileGluLeu-----GlyLeuAsnPheLysAsnGluAlaGln---GlnLysTh 1728
    ::::::::::::::::::::
DB 6029 CCAGAGAGTGAAGAACTTCTGAAGGCTCAATTCGATTAGAAATGATGCAATGAATTC 6088
OY 1728 rThrLysGluGlnCysleuLeuAsnLys-----1739
    ::::::::::::::::::::
DB 6089 ATACCGGAAGATATGAGATTAATGTGGCCAAAGTGAATGACAGCTGAGAGAGATTT 6148
OY 1740 ---GluLeuGlnGlnSerGlnHisArgleuGlnCysGlu-----1751
    ::::::::::::::::::::
DB 6149 TCTTGATGTGAAAAATGAGCTGAGTGAATCAGATCGGAAACCTGACATGAGCATGA 6208
OY 1752 -----lleGluGluLeuMetLysSerleuLysAspLysGluSerAlaLeuGluTh 1768
    ::::::::::::::::::::
DB 6209 AGCCCTTACTGTGAGGCTGACTTAGAGTATTCAAACAGAAACCTATGTTAGAAAA 6268
OY 1768 rLeuLysGluSerGlnLysValIle---AsnleuAsnGlnGluMetGluMetValMe 1787
    ::::::::::::::::::::
DB 6269 AGACAATGAATAATGACGAGAGATTAATGTCTGCTGAAGAAAGCTCTCAGTGTCA 6328
OY 1787 tLeuGluMetGluLysleuLysAsnSerGlnArgThrValIleAlaGluArg-----1804
    ::::::::::::::::::::
DB 6329 AAGTGAAGAAACCGCTGTGTGAGAAATTAATGATTAATCAAAAAAACCAGCGCACT 6388
OY 1805 -AspGlnleuGlnAspAspLeuArgGluSerValGluMetSerlleGluThr-----1821
    ::::::::::::::::::::
DB 6389 GGATCAGTGTCTGAAAAATTAAGAGAAACACAA---GACCTTGAGTGTATCAAAAG 6445
OY 1822 -----GlnAspAspLeuArgLysAlaGlnGluAla 1832
    ::::::::::::::::::::
DB 6446 TGAGTGTCTCATTCATTCAGTGCAGAGGCGAGAGGTGAAGAAAGACGAACTCCT 6505
OY 1832 uGlnGlnGlnLysAspLysValGlnGluLeuThrSerGlnIleSerValleuGlnGluLys 1852
    ::::::::::::::::::::

```


Db 6506 TCAGACTTTGCTCTGATGTAGTGTAAAGACAAACATCATCTCCAGAGAA 6565
 QY 1852 sIlleSerLeuLeuGluAsnGlnMetLeuYrsnValAlaThrValLysGluThrLeu 1872
 Db 6566 GGTGCAGACTTTGGAAAAGAGACTCACAGGCACTCTTTGCAAAAATGTGAGCTGAGAAA 6625
 QY 1872 rGluArgAspArgLeuAsnGlnSerLysGlnHisLeuPheSerGluIleGluThrLeu 1892
 Db 6626 CCAAAATGCAACAACGTAATTAAGAGAAAGAAATTCCTTGTTCAGAGATTCGAAAGCTTGA 6685
 QY 1892 rLeuSerLeuLysGluLysGluPhe-----AlaLeuGlnAla 1905
 Db 6686 GCCACAGACTGAGTATCATCTATGAAAAGCTGAATGTCTCCAGCCCTTGAGGCGCC 6745
 QY 1905 aGluLysAspLysAlaAspAlaAlaArgLysThrIleAspIleThrGluLysLeuSer 1925
 Db 6746 ACTGTTGAGAAAAGGTGAGTTCGATGTAGGCTGAGCTCACACAGAGAGAGAGTATCA 6805
 QY 1925 nIle-----GluGlnGlnLeu 1931
 Db 6806 GCTGAGAAAGAGCATCGAAGAACTGAGACTTCGATTCAGCCGATGAAAGAGAGCT 6865
 QY 1931 uGlnGlnAlaThrAsnLeuLysGluThrLeuYrGluArgLysSerLeuIleGlnCysLy 1951
 Db 6866 GCACATCGCAGAGAAACTGAAAGAACGCGAGCGGAGATGATTCCT-----AA 6916
 QY 1951 aGluGlnLeuAlaLeuAsnThrGlnHisLeuArgLysThrLeuLysSerLysAspLeu 1971
 Db 6917 GCATTAAGTTGGAACCTTGAAAGGAAATTCAGATGTCAAGAAAGACAGAGCTAGT 6976
 QY 1971 aLeuLysMetGluGlnGlnLysArgAspGluAlaAlaAsnLysValIleAlaLeuThrG 1991
 Db 6977 GAT-----CTTGATGCCGAGAAATTCAGACA----- 7004
 QY 1991 uLysMetSerSerLeuGluGlnGlnLeuAsnValIThrThrLeuLysGluLys 2011
 Db 7005 -GAAATGAGACTGTAAAAACAATATAGAGAGAGCGCAGAGCTTAAAGTTTGA 7063
 QY 2011 u-----GlyLysGluThrPheYrLysGlnArgProSer 2024
 Db 7064 ATTACACCTGTACGTTAAGTCTGAAAAAGAAAATCTGACAAAACAATACAGAGAA 7123
 QY 2024 sGlnGlnSerSerGlnMetGluLysGlnSerLysSerLeuLysAspLeu 2044
 Db 7124 ACAAGCTCAGTTGAGAACTAGACAACTTCTCTCATTTAAAGCTGTAGACA 7183
 QY 2044 nLeuGluAlaGluLysGluLysSerGluAlaThr----- 2056
 Db 7184 AAAGAGACAGAGATACAGATCAAAAGAAATCTAAACTGCAGTGAATGCTTCA 7243
 QY 2057 -AsnGluLysAsnLeuThrAlaLysIleSerSerLeu-----GluGluGluLe 2074
 Db 7244 GAAATAGTTAAAGAGCTTAATAGAGCACTACAGCTTGTGTGTGTCAGAAATTA 7303
 QY 2074 uGln---AsnAlaSerIleLeuAsnGluAlaValSerGluArgLysAspLeuArgHis 2093
 Db 7304 GAAGCCACAGAGAGAGCTAGACCAACCAATAGAGAGAGATCAGCTAGAAATAG 7363
 QY 2093 rLysGlnGlnLeuValSerGluLeuLys-----GlnLeuSerLeuThrLe 2108
 Db 7364 CATTAAAGAGCTGAGAGCCCGCTAGAGCTGATGAAAAGAAAGCTTGTCTTACA 7423
 QY 2108 uLysSerArgAspHisAlaPheAlaGlnSerLysArgLysLysAspGluAlaValAsn 2128
 Db 7424 ACAACTGAG-----GAAGTGAGCATCTGACAGATTTACTTAAGGTAG 7468
 QY 2128 sIleAlaSerLeuAlaGluLysIleLeuThrLysGluMetAsp----- 2144
 Db 7469 AGTGAGAGACCTTGAAGAGAGCTAGAGTACCAAGAGAAACAAGACATGACAGCT 7528
 QY 2145 -GluPheArgAspSerLys-----GluSerLeuGlnGlnSerSerHisLeu 2161
 Db 7529 TGAGCAGAGAAATTCAAAAGAGAGTAGAGACCTTAAAGCAAAAATAGAGAGATGAC 7588

QY 2161 rGluGluLeuCysThrTyrLysThrGluLeuGlnMetLeuLysGlnLysGluAsp 2181
 Db 7589 CCAAGCTGTAGAGCTGTGGAAATAGATGTGTACTATTAAGGTGAGAAAAGAAATCT 7648
 QY 2181 eAsnAsnLysLeuAlaGluLysValLysGluValAspGluLeuGlnHisLeuSer 2201
 Db 7649 GCAAAATGATATCAAAAAGAGCAAGAGCAAAATATCTGAATTA----- 7691
 QY 2201 rLeuLysGluGlnLeuAspGlnIleGlnMetGluLeuArgAsnGlnLysLeuArgAsn 2221
 Db 7692 -----GAAATTAATAATTCATCATTTGAAAT-- 7718
 QY 2221 rGluLeuCysGluLysMetAspIleMetGluLysGluLysSerValLeuArgLeuMet 2241
 Db 7719 -----ATTTCCAAAGAAAAGA 7735
 QY 2241 nAsnGluProGlnGlnGluLysAspValAlaGluArgMetAspIleLeuGluSerArg 2261
 Db 7736 GCAAGAGAAAGTACAGATGAAAGAAAATCAACGCACTGCCATGAGATGCTTCAACA 7793
 QY 2261 gAsnGlnLysIleGlnGluMetGluLysIleSerAlaValTyrSerGluGlnHis 2281
 Db 7794 -----CAATTAAGAGCTCAATGAGAGAGTGGAGCGCTGCAATATGACCA 7841
 QY 2281 rLeuLeuSerSerLeuSerSerGluLeuGlnLysGluThrGlnAlaHisLeuHis 2301
 Db 7842 -----GAAGCCTGAG----- 7853
 QY 2301 rLeuAsnIleLysGluSerLeuSerSerThrLeuSerArgSerPheGlySerLeuGln 2321
 Db 7854 -----GCCAAGAGCAGATAGTACTAGTCAAGTAGAGTCTTGAAC 7897
 QY 2321 rGluHisValLysLeuAsnThrGlnLeuGlnThrLeuLeuAsnLysPheLysVal 2341
 Db 7898 TGAGAAAGCTCAGTTGCTCAAGCGCTTGATGAGGCCAAAATTAATTAATGTTTGA 7957
 QY 2341 rArgThrAlaAlaValLysGluAspHisSerLeuIleLysAspTyrGluLysAspLeu 2361
 Db 7958 ATCTTCAGG-----AATGCCATTCAGAAAGTGAAGATGCCAGCA 8002
 QY 2361 aAlaGluLysArgHisAspLeu-----LeuArgLeuGlnLeuGlnCysLeu 2378
 Db 8003 GAAATCGAGAAAGAGAGAGAAATCAGTAGACTGAGAAAATCAAAATTCAGAGCAGA 8062
 QY 2378 uGlnHisGluArgLysTrpSerAspSerAlaSerGluGluLeuLysPhe----- 2394
 Db 8063 GCAGCTTGTCTTAACTGTCCAGGTGAGAGAGAGACCAACTTGGAGAGAGCAAAA 8122
 QY 2395 -CysGluIleGluPheLeuAsnGluLeuLeuPheLysLysAlaAsnIleIleGlnSer 2414
 Db 8123 CTTAGACAGAGAAATCTACAGTGAATGGAGCGAAAGATCCAAAGTCTCAATTC 8182
 QY 2414 lGlnAspAspPheSerGluValGlnValPheLeuAsnGlnValGlySerThrLeuGln 2434
 Db 8183 AATGCTCTTTCAGAGACACATTAGAGTGTGAGAGTCTTACAAAGATCTAGAGAA 8242
 QY 2434 uGluLeuGlu---HisLysLysGlyPheMetGlnTrpLeuGluGluPheGly----- 2450
 Db 8243 TGAGCTGAATTCACAAAATGAGCAAAATGTCCTTGTGAAAAGTAAACAAATGAC 8302
 QY 2451 -----AspLeuHisValAspAlaLysLys-----Le 2459
 Db 8303 TCAGAAAGAAAGTACAGCTCAGAGGAAATTCATGAGATGCGACAGAAAACAGAGCT 8362
 QY 2459 uSerGluLysMetGlnGlnGluAsnArgLysIleAlaSerThrIleGlnLeuLeuThr 2479
 Db 8363 GCAAGAGAAACACAGTGAAGAGAAAATAGCGTGAAGAGTTCAGTCTTGTGA 8422
 QY 2479 sArgLeuLysAlaValValGlnSerLysIleGlnArgGluIleThrValTyrLeuAsn 2499
 Db 8423 AGAAATTAAGAGC-----AGCAAAAGATCA----- 8447

QY 2499 nphleuAlaLysLeuIngluLysGluInaLysGluLeuMetArgArgMetGI 2519
 Db 8448 -----TTGAGAGGCTCACACTAGAAAATGTAATGGAAG----- 8486
 QY 2519 uHISHisGlyProSerAlaSerValMetGluGluValAsnAlaArgLeuLeuGlyIleLe 2539
 Db 8487 -----AGCTTACATTCATGCACAAAGACAGCTGGAAGAGAGG----- 8528
 QY 2539 uLysThrValGlnAspGluSerLysLysLeuGlnSerArgIleLysMetLeuGluAsnGI 2559
 Db 8529 -----AAAGTGAGAGAGAAATAGCTGTAATATCAGCTACGG----- 8564
 QY 2559 uLeuAsnLeuValLysAspAlaMetHisLysGlyGluLys-----ValAlaIleLe 2577
 Db 8565 -----CTTCATGAGCTGAAAAGAAACACACAGGCTTGCT 8599
 QY 2577 uGlnAspLysLeuLeuSerArgAsnAlaGluAlaGluLeuAsnAlaMetGlnValLysLe 2597
 Db 8600 TTTCGAC-----ACAACAAACAGTATGAAAGTAATCCAGACATACCGAGAAATTT 8653
 QY 2597 uThrLysLysGlnAspAsnLeuGlnAlaAlaMetLysGluIleGluAsnLeuGlnLysMe 2617
 Db 8654 GACTTCTAAAGAAAGAAAGTCTCAGTCAACAGCTGAGATAGACCTTTTAAAGTCTAG 8713
 QY 2617 tValAlaLysGlyAlaValProTyrLysGluGluIleAspAsn----- 2631
 Db 8714 T-----AAAGAAAGCTCAATTAATTCATTTGAAGCTACTAC 8749
 QY 2632 -----LeuLysThrLysValValLysIleGluMetGluLysIleLysIyr----- 2666
 Db 8750 TCAGATTTTGGAGAAATGTAAGAAACCAACAGTGCACATCTAAATATGTAATCAGTT 8809
 QY 2647 -----SerLysAlaThrAspGlnGluIleAlaLysSerCys----- 2660
 Db 8810 GAAGAGAAATATGACCTGCCAGGGGAAATGATGCTTTCATCAATCTCTAACA 8869
 QY 2661 -LeuGlnAspLysGluGluGluLysArgLysLeuLysGluLysLeuArgArgAlaGlnAl 2680
 Db 8870 GCTGGAAGAGAAAGAGATCTGCAGAAAGAACTCTCTCAACTCAAGCTGCACAGGA 8929
 QY 2680 aAspAsnAspThrThrValLysValProLysAspTyrGlnLysAlaSerThrProProVa 2700
 Db 8930 GAAGCAGAAACA----- 8942
 QY 2700 lThrCysGlyLysGlySerGlyIleValGlnSerThrAlaMetLeuValLeuGlnSerGI 2720
 Db 8943 -----GGTACTGTATGATATCAAGGTCGATTAACAACACTGA 8983
 QY 2720 u-----LysAlaAlaLeuGlu-----ArgLysLeuSerHisLysIyrLys 2733
 Db 8984 GATCAAGAACTGAAAGAACTCTTGAGAAACCAAGAGGAGCATGTAATCTTGA 9043
 QY 2733 sLysThrHisHisLeuSerArgThrMetSerSerSerGluAspArgLysLysThrLysAl 2753
 Db 9044 TAAGTACGCTTCCTG-----CTTATTAAGCATTAAGAAAGTATAGAAAGCTTAAGA 9094
 QY 2753 aLysSerAspAlaHisSerSerHisThrGlySerSerHis-----ArgGI 2768
 Db 9095 GATGTTAGAGACACAAGTGGCCACTGTCTACAGCAATCTAAACAAGATTCGCCGAG 9154
 QY 2768 ySerProHisLysThrGluThrThrArgHisGlyProValThrProGluArgSerGluMe 2788
 Db 9155 GTCTCTCT-----TTGCTAGTTCATGTTCTTCAGAGACATCTCCAT 9196
 QY 2788 tProSerLeu-----HisLeuGlySerProLysLysSerGluSerSerThrLysAr 2805
 Db 9197 CCTTCTCTTACTGAAGAGGTTATCATCTGCCAAATATAAGCTTCAGCAAGAGCA 9256
 QY 2805 gValAlaSer-----ProAsnArgSerGluIleLety 2815
 Db 9257 AAGATCCAGTGGAAATATGGAGCAATGTGGAGACCAACACCTGCTTACCCAGAGAGCTT 9316
 QY 2815 rSer-----GlnLeuValMetSer-----ProGlyLys----- 2824

Db 9317 TTTCTAAAAAGCAAGAAAGCAGTATAGTGTATTCACCTGCAGAAACACGGAAG 9376
 QY 2824 ----- 2824
 Db 9377 TACTGATTTGAGCCAGAGGACTTCACAGACTTGAAGAAAGGTTTCTGCATCC 9436
 QY 2825 -ThrGlyMetHisLys-HisIle-----LeuSerProSerL 2836
 Db 9437 GACAGGAAAGACTAGCCCTATATCTCTGCGAAGAACCAACCATGGCACTCGGACAGCC 9496
 QY 2836 yValAlaGlyLeuHisLysLysArgAlaLeuSerPro----- 2847
 Db 9497 CCAGCTGCTGCACAG-AAAGTTAGCCCTATCCCACTGAGTCTCGGCAAGAAATCTTG 9555
 QY 2848 -----AsnArgSerGluMetProThrGlnHisValI 2858
 Db 9556 CAGAGTCTCCAAACCAACAGCTGTGGCAGCAGATCAACAAAGTCATGCTCAGC 9615
 QY 2858 LeSerProGlyLysThrGly-----LeuHisLysAsnLeuThrGluSerThrLeuPhea 2876
 Db 9616 GGAGCCCAAGATATCAGGACCATCTCCAGAAACCCACCAAAATCCGTCCAGTCA 9675
 QY 2876 sPasnLeuSerSerProCysLysGlnGluValGlnGluAsnLeuAsnSerProLysG 2896
 Db 9676 ATATATCTCTGAGAGAAAGTCCAGTGCACAGCCCAAGAGGGGCTGAGGCTCAGCG 9735
 QY 2896 lLysLeuPhe-----AspValLysSerLysSerMetProTyrCysPro 2911
 Db 9736 GCCGACTTTCGCCAGCCCAAGCTGAGCTGAGTCAAGGCAATGAGAACTGTAAAG 9795
 QY 2911 eArgL 2912
 Db 9796 TCCAG 9800

RESULT 5

US-09-864-864-292
 ; Sequence 292, Application US/09864864
 ; Patent No. US20020102679A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Mitcham, Jennifer L.
 ; APPLICANT: Harlocker, Susan L.
 ; APPLICANT: Dillon, Davin C.
 ; APPLICANT: Secrist, Heather
 ; APPLICANT: Iodes, Michael J.
 ; APPLICANT: Fling, Steve P.
 ; APPLICANT: Benson, Jane R.
 ; APPLICANT: Carter, Darick
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; FILE REFERENCE: 210121.523
 ; CURRENT APPLICATION NUMBER: US/09/864, 864
 ; CURRENT FILING DATE: 2001-05-23
 ; NUMBER OF SEQ ID NOS: 341
 ; SOFTWARE: Corixa Invention Disclosure Database
 ; SEQ ID NO 292
 ; LENGTH: 10190
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-864-864-292

Alignment Scores:

Pred. No.: 7.97e-64 Length: 10190
 Score: 1139.00 Matches: 703
 Percent Similarity: 37.36% Conservative: 556
 Best Local Similarity: 20.86% Mismatches: 1052
 Query Match: 7.71% Indels: 1060
 DB: 10 Gaps: 139

US-09-150-867-1 (1-2954) x US-09-864-864-292 (1-10190)

```

OY  * 333 ThrProHisValAsnGluValLeuAspArgIleAlaLeuLeuLysArgTyrGlySer 352
DB 112 ACAGAGCTCTCAGAAAAATTCAGAGCTTGAGAGCAGCTTGACAACTGAAAGAGAA 171
OY 333 IleuAspLeuLysGluLeuGluAsnLeuGluSerSerGluThrLysAlaGln 372
DB 172 AAGCAGCAAGAGCAGTTTCAGCTTGACAGCTCGAGCGCTGCAGAAACAGACAG 231
OY 373 AlaMetAlaLysGluGlu-----HisThrGlnLeuLeuAla 384
DB 232 AAGGTGAAATGAAAAAACCGAGGTCACAACTGAAAAAGGAGAAATCAAGATTGAAG 291
OY 385 GluIleLysGlnLeuHisArgGluArgLysArgIleThrHisLeuThrAsnIleVal 404
DB 232 GAAATATGTGAAGAGCTGAGAGAAACTAAGCAGAAATTTCTCAT-----GAACTTCAA 345
OY 405 ValAlaSerSerGlnGluSerGlnGlnAsp-----GlnArg 416
DB 346 GTCAAGAGAGTCAAGATGAAATTTCCAGAGAAAGCAACTGAAATTCAGGCAAAAAACAATA 405
OY 417 ValLysArgLysArgArgValThrThrAlaProGlyLysIleGlnAsnSerIleuHisAla 436
DB 406 GAAAAAGTGGAAACAGAACTTAAAGGTGTAAATCTGACCTGAAAGAACCCAGACAGCT 465
OY 437 SerGlyValSerAspPheAspMet-----Leu 445
DB 466 GCGCAGTGTGCGAGATGTCTCTGTGAATCCATGCATACACCAAAAAATTTTACAACT 525
OY 446 SerArgLysProGlyAsnPheSerLysLysAlaLysPheSerAspMetProSer----- 463
DB 526 CCACATACACCAAGTCATATATATAGTGTCCAAAGTATGAAAGCTAAAGAAAAAATAT 585
OY 464 PheProGluIleAspAspSer-----ValCysThrGluPheSer 476
DB 586 AATAAAGAGGTGAAAGAAAGAAAAAGATTAGAGCAGAGGTTAAAGCTTGAGAGCTTAA 645
OY 477 AspPheAspAlaLeuSerMetLeuAspSerAsnGlyIleAspAlaGluThrAsnLeu 496
DB 646 AAAGCAACCCAGACCTTCCACAAAGCCACAGCATGATCACCAGCATTTGCCGCGATCAG 705
OY 497 AlaSerLysValThr-----HisArgGluLysThrSerLeuHis----- 509
DB 706 GCTTCATCATCTGTGTCTCATGCGACAGAGAGAACCCCAAGCATCTTTCATCATAT 765
OY 509 ----- 509
DB 766 TCTCAAGAACTCCAATTAGAGAGATTCTCTGCATCTTACTTTCTGGGAGACAGAG 825
OY 510 -----GlnSerMetIleAspPheGlyGlnIleSerAspSerValGlnPheHis 525
DB 826 GTGACTCCAGATCGATCACTTTCGCAATAGGAAAAAGAGATGCTAATAGCGATTCTTT 885
OY 526 AspSerSerLysGluAsnGlnLeuGlnIleuProLysAspSerGlyAspMetAlaGlu 545
DB 886 GACAATCTTACAGACTCTCATCTTTG-----GATCAATTAATAA 924
OY 546 CysArgLysAlaSerPheGluLysGluIleThrSerLeuGlnGlnGlnLeuGlnSerLys 565
DB 925 GCGCAGAAATCAAGAGCTTAAAGAAACAAAGATTATAGATTGGAACCTGAGCTGCAAGACAT 984
OY 566 GluGlnGluLysLysGluLeuValGlnSerPheGlnLeuLysIleAlaGluLeuGlnGlu 585
DB 985 GAAAAAGAAATGAAGAGCCAGCTGATTAAGTTTCA-----GAACTCCAACTC 1032
OY 586 GlnLeuSerValLysAlaLys-----AsnLeuGluMetValThrAsnSer--- 600
DB 1033 CAACCTGGAG-----AAAGCAAAAGTGAATTAATTGAAAAAGAAATTGTTGAACAAATGT 1089
OY 601 ArgGlnHisSerIleAsnAlaGluValGlnThrAspValGluLysGluValValArgLys 620
DB 1090 AGGGATGAAGTATGAGAAACAAACAGCAATATGAC----- 1125

```

```

OY 621 GluMetSerValLeuGlnLysAspSerGlyTyrAsnAlaSerAsnSerAspLeuGlnAspSer 640
DB 1126 -----CAGCGCTCAACCAAGTATACTGCTTGGAAACMAAAATGAAAAATTTG 1173
OY 641 SerValAspGlyLysArgLeuSerSerHisAspGluCysIleGluHisArgLysMet 660
DB 1174 ACGGAAGAT-----TTGAGTTGTCAGCGCAGCAAAATATGCAAGAAAGTCCAGATGTTCT 1224
OY 661 LeuGlnGlnLysIleValAspLeuGlu-----GluPheIleGluAsnLeuAsnLysLysSer 679
DB 1225 CTGGAACAGAAAAATTAAAGAAAAAGAGATTTCAAAGAGAGCTTCCTCCCTG----- 1278
OY 680 GluAsnAspLysGlnLysSerSerGlnGlnAspPheMetGluSerIleGlnLeuCysGlu 699
DB 1279 -----CAACAGCGTCTTTCCAACACCTGACAGCAGAGATGCTCCAGATG----- 1323
OY 700 AlaIleMetAlaGluLysAlaAsnAlaLeuGlnGluLeuAlaLeuMetArgAspAsnPhe 719
DB 1324 -----AAGCCAGACTCACCCAGAGATTACAGCAAGCCAAAG-----AATATG 1365
OY 720 AspAsnIleIle-----LeuGluAsn 726
DB 1366 CACAAAGCTCTCAGGCTGAACTGATTAATCACATCAGTAAGCAACAGCAGCTAGAAAC 1425
OY 727 -----GluThrLeuLysArgGluIleAlaAspLeuArgSer----- 739
DB 1426 AATTGGAAAGATTATAGCAAAAGTTGTCCAGACGTCAGAAACAGCGCTTCCAGCGAGTCA 1485
OY 740 LeuLysGluAsnGlnGluThrAsnGluPheGluIleLeuGlnLysGluThrGln----- 757
DB 1486 ATCAAGAGAAATGAGCTGAGAGAAAGATGAGAAATGAAAGAAACAAACCTCTT 1545
OY 758 LysGlnHisGluAlaGlnLeuIleHisGluIleGlySerLeuLysLeuValGluAsn 777
DB 1546 AAGAGTCACTGTGAGCAAAAGGCCAGAGAAAGTGTCCACTGTGAGGACAGAACTCAAGAC 1605
OY 778 AlaGluMet-----TyrAsnGlnAsnLeuGlnLysAspLeuGlnThrLysThrLys 794
DB 1606 ATCAAAACAGTGTTAATTCAGACGAATTTTGCAGAAAGAAATGAAACCGAAGAT----- 1662
OY 795 LeuLeuLysGluGlnGluIleGlnLeuAlaGluLeuArgLysArgAlaAspAsnLeuGln 814
DB 1663 -----ACCTTCAGAGAAACATGTTAAGAGATCTTCAAGAAATAATATCAGCAAGAA 1716
OY 815 LysLysValArgAsnPheAspLeuSerValSerMetGlyAspSerGluLysLeu----- 832
DB 1717 AACTCCTTCACTTAGAAAAAACTGAAGCTTGTGTGTGCGATGTGAAAAACGACGAGAT 1776
OY 833 CysGluGlu-----IlePheGlnLeuLysGlnSerLeu 843
DB 1777 TCTTCTCAAGACCTTTTGAAGAAAAAGAAACATCACATTTGAACAACTTAAATGATTAAGTTA 1836
OY 844 SerAspAlaGluAlaValThrArgAspAlaGlnLysGluCysSerPheLeuArgSerGlu 863
DB 1837 ACCAAGACAGAGAAAGAGTCCAA-----GCCCTGTGAGTGTGAGTTAAAGAAAGAA 1893
OY 864 AsnLeuGlnLeuLysGluLysMetGluAspThrSerAsnThrTyrAsnGlnLysGluLys 883
DB 1894 GAATATGAATTTGAAAGAGAAAACTGTGTTTGTGTTGAAAGAGTCAAAAAACGAAAAA 1953
OY 884 AlaAlaSerLeuPheGluLysGlnLeuGlnThrGluLysSerAsn----- 898
DB 1954 CTTTAACT-----CAGATGGAATACAGAAAAAGAAACTTTCAGAGTAAATTT 2001
OY 899 -----TyrLysLysMet 902
DB 2002 AATCACTTGGAAATCTGTCTGAAGACACAGCAAAATTAAGTCAATGATACAAACAGAGA 2061
OY 903 GluAlaAspLeuGlnLysGluLeuGlnSerAlaPheAsnGluIleAsnTyrLeuAsnGly 922
DB 2062 GTAAAGACGCTGAGATGAGACAGAGAAACCTTAAGTGTGAGATCAGAAACCTTCAACAC 2121
OY 923 LeuLeuAlaGlyLys-----ValProArgAspLeuLeuSerArgValGluLeuGln 939

```

Db 2122 GTGTTACACAGTACAGTAGAGAGACCCAGAACTACTATATAGAGCTACAG 2181
QY 940 LysLysValSerGluPheSerLysGlnLeuGlnLysAlaLeuGlnLysAsnAlaLeu 959
Db 2182 CAGAAA--GCTGCTTCTCA-----GATCGAAACATCGAAG 2217
QY 960 GluAsnGlnValThrCysLeuSerGluThrLysPheLeuProAsnGlnValGluCysLeu 979
Db 2218 GAATATGAAAATATATGTGTTGAAAGCTTCACG--CTTACTGGCGAAGTTGAAGATCTA 2274
QY 980 LysAsnGlnLysSerLysAlaSerGluGlnLysLeu----- 991
Db 2275 GAACACAAAGCTTCACTACTCTCAATGAATATGACAAAGACCGGTGTACCAAGAC 2334
QY 992 -----LeuLeuLysGlnGluLys----- 997
Db 2335 TTGCATGCCAATATGAGAGCCTCAGGAGTCTGTAATAATCCAAAGATGCTCTGTGTG 2394
QY 998 -----GluHis 999
Db 2395 ACAATGAGATCATCAGAGAGTCTTTGGCTTTGATCAGACGCTGCCATGCATCAT 2454
QY 1000 Ser--AlaSerIleIleSerLysGlnLysLeu----- 1010
Db 2455 TCCTTTGCAATATATATGGAACACAGAGCATGCCCTTCAGAGAGAGTGAATGTGT 2514
QY 1010 ----- 1010
Db 2515 TTAGAACGACAAAGTCCGAAATTCCTCCATCTTCACAAATAGATTGATCACT 2574
QY 1011 -----MetGlnGlnInsSerGluGln 1017
Db 2575 GAATTTTCATAGACTCTCAAAAACAGATGACATCAGACCTGCGAAACGATGGAAGAG 2634
QY 1018 IleLeuGlnLeuThrAspGlnValThrHisThrGlnSerLysValGlnGlnThrGluGln 1037
Db 2635 TTGGTGCAAATCAAGAGAAATAGAAATAATCATGAAAGCAAGCAGATGCATCA 2694
QY 1038 GlnThrLeu-----GlnMetLysMetHisAspAsp----- 1048
Db 2695 AGTTTGTGGCTGAAACAGATCAGCGCATTAAGTTACAGAGAACATCTTCCTCAC 2734
QY 1048 ----- 1048
Db 2755 CAGAAATGTTGTCTGAAACCTTAAGTCCCTTGAGAACAGAAAGAGCTGCACAT 2814
QY 1049 LeuPheGlnLysThrIleArgAsnLysSerGlu----- 1059
Db 2815 TTAAATGATAGAGTAAAGACTGACGAGCGAGATTCAAGAAATTAAAAAGACACAT 2874
QY 1060 ---AlaGluAspLeuLeuArgGlnMetGlnAsnLeuLysGlnThrMetGlnSerValGln 1078
Db 2875 CTACTTGAAGACTCTCTAAGAGAGCTACACATTTATCGGAACCTTA--AGCTTGAG 2931
QY 1079 ValLys-----IleAlaAspThrLysHisGlnLeuGlnGluThrIleArg 1093
Db 2932 AAGAAAGAAATGAGTTCATCTCTTAATTAAGGAAATGAAAGAGTGCACCA 2991
QY 1094 Asp-----LysGln-----GlnLeuLeuHisGlnLysThrPhePhe 1106
Db 2992 GAGAAATGAGCTCTTAAGGAAATTAATGCATCTTAATCAAGAGAGGAACTTAATC 3051
QY 1107 GlnAlaMetGlnThrIlePhe-----ProIleThrProLeu 1118
Db 3052 CAGAAAAGTGAAGTTCGCAACACTATATAGATGAAGGAGGAAAGCAATTCACAGTTA 3111
QY 1119 SerAspSerLeuProSerLysLeuVal-----GluGlnAsn 1131
Db 3112 TCCTGATCAGTACAGCAAGAAAACCTATTATTACTACAAAGATGGAAGAAACCGGAAT 3171
QY 1132 SerGlnAspProIle----- 1136

Db 3172 GCATATGAGGATCTTAGTCAAAATAATACAAAGACGACAGAAAAGATTTCTAAATGAA 3231
QY 1136 ----- 1136
Db 3232 TCTCTGTAATGAAATGCATAGTCTTTGTGAAATATAGAAAAATGACTGGACAGCTA 3291
QY 1137 -----GluIleAsnAspThrHisAsnLeuIleAlaLeuAlaThrGln 1150
Db 3292 AAGGAAGCATTTGCAAGAGAACACCAAGAAATCTTAACAAAATTAGCATTTGCTGAGAA 3351
QY 1151 ArgAsn--AsnIleMetValCysLeuGlnThr----- 1160
Db 3352 AAGAAATCAGAAATCTGATGCTAGATGGAGACAGTGCAGCAAGCTCTGATCTGAGATG 3411
QY 1160 ----- 1160
Db 3412 ACAGATTAACCAAAACATTTCTAAGCGAGGTGGGTGTTAAGCAAGAAATCATGACT 3471
QY 1161 -----GluArgAsnSerLeuLysGlnGlnValIleAspLeu-----AsnThr 1174
Db 3472 TTAAAGCAAGAACAAACAAATGCAAAAGAAATTAATGACTTTATTCACAGAGATGAA 3531
QY 1175 GlnLeu-----GlnSerLeuGlnAlaGlnSerIle 1184
Db 3532 CAGCTGATGAAGGTAAATGAAAGACTTAACATGATGTCAAAATCTGAATCAGAACCAT 3591
QY 1185 GlnLysSerAspLeuGln-----LysProLysGln 1194
Db 3592 AGGACTCTCTGAAAGAAAGAGAGAGTAGAGAAATCAATTAATTTTAACTCAGATG 3651
QY 1195 AspLeuGlnGlnLysGlnValLysLeu-----LeuLeuGlnMetGln 1208
Db 3652 GATCTTGAACTTAAAGAAATTTCTAGATAGTATATATGCGAGTGGTGCAGAAATGAA 3711
QY 1209 LeuLeuLysGlnHisLeuThrAspSerGlnLeuSerIleGlnLysLeuGlnLeuGln 1228
Db 3712 GCATG-----CTAAGAAATAGGAATTAACCTCAGAGAAAGTGAAGAGAGAG 3762
QY 1229 LeuGlnValThrGlnLysLeuGlnThrLeuGlnGlnGlnMetLysAsnIleThrIleGln 1248
Db 3763 -----GAGTGCCTGCACAGATGAATTAAGCAAT----- 3792
QY 1249 ArgAsnGlnLeuGlnThr--AsnPheGlnAspLeuLysAlaGlnHisAspSer--Leu 1266
Db 3793 AGAGGAGATCTTGAACACACACATTTTGCAAGACATGACAGACAAAGAAATTAAGTGCCT 3852
QY 1267 Lys-----GlnAspLeuSerGln 1272
Db 3853 AAAGACTGTGAATAGATGCGGAAAGAAAGTATATTCAGGGCTCATGAGTTGTCAACA 3912
QY 1273 AsnIleGlnInsSerIleGlnThrGlnAspGlnLeuArgAlaIleGlnGlnGlnLeuArg 1292
Db 3913 AGTAAAGACGACAAATGACACACCTTCACTGCTCTGCAACCAACATGAAACAGCGAAT 3972
QY 1293 GlnGlnLysGlnLeuValAspSerPheArgGlnGlnLeuLeuAspCysSerValGlyIle 1312
Db 3973 GAGCTAGAGAAATATGTGAATATCTCAGGCGTGAAGAAAGTATGACCTGACTGAGCTG 4032
QY 1313 SerSerProAsnHisAspAlaValAlaAsnGlnLysValSer----- 1327
Db 4033 AATGATTCAGAGTCAAGATGATACACAGCAACTAGGAAATGCGACAGAGAGTAGGAAA 4092
QY 1328 ---LeuGlnGlnValAsnSerLeu-----GlnSerGlnMetLeuThrGlnGlnValAsp 1344
Db 4093 CTACTTAATGAAAGTAAATTAATTAATGATGATGACAGTGTCTTCCATGCTGAGTTAGTG 4152
QY 1345 Gln----- 1345
Db 4153 GAAAGCATACAGAGAGTGAATTTGGTGAACCAACCAATGAACAGCACACCTGTGTCTTG 4212
QY 1345 ----- 1345
Db 4213 GCTCATTTGACGAGAGATATTCCTAGAGACATTTGACATTTGTCAGACAAAGATTCAA 4272

QY 1346 -----LeuGlnThrSerCysLysAlaLeu 1353
 Db 4273 ATGCACCTTGGCGAATTGCAAGAGAAATTTCTTACTTTCACAAAGGAGAACACAAATTTTAA 4332
 QY 1354 -----ValSerGluLeuGluLeuLeuArgAlaHisValLysSer 1366
 Db 4333 CATGATCAGCACTGTACAGATGAGCTCTAAATGCTCAGAGCTGACAGACCTATGTTGACTCA 4392
 QY 1367 ValGluGluAsnLeuGluIleThrLysLysLeuAsnGlyLeuGluLysGluIleLeu 1386
 Db 4393 TTTAAAGCCGCAAAATTTGGTCTGTTCACAGATCTGAGAAAATTTCAGAAAGTGACTGTGCTG 4452
 QY 1387 GlyLysSerGlu-----GluSerGluValLeuLysSerMetLeuGluAsnLeuLys 1403
 Db 4453 AAGAGATGCAAGCTGGGCTGTGGAGAGGGGCTGCTCCATCTCCCTGTCATCTCTTGTGTG 4512
 QY 1404 GluAspAsnAsnLysLeuLysGluGlnAlaGlu----- 1414
 Db 4513 CCTGACAGCTCTAGCTTAGCAGTTTGGAGACTCCCTTTTACAGAGCTCTTTAGAA 4572
 QY 1415 -----GluTyrSerSerLysGluAsnGlnPhe 1423
 Db 4573 CAGACAGAGATATGCTCTTTTGAATATTAGAAAGGGGCTGTTTCAGAAACCCAGTGC 4632
 QY 1424 SerLeuGluGluValPhe----- 1429
 Db 4633 AGTGTAGATGAATATTTTGCAAGCACTGTGACAGAGAGAACTGACAGAGAAACCC 4692
 QY 1430 ---SerLysSerGlnLysLeuValAspGluLe-----GluValLeuLysAla 1444
 Db 4693 CCTTCGGCCCCAGAGGGGTGTGAAGAGCTTGAGTCCCTGTGAGAGCTGTACCCGAG 4752
 QY 1445 GlnLeuLysAlaAlaGluGluArgLeuGlu-----IleLysAspArgAspTyr 1460
 Db 4753 TCCTCGAGAACTAGAGAGAGAAATGAAAGTCAAGGATTAAGAAATAGAGAAATTT 4812
 QY 1461 PheGluLeuValGlnThrAlaAsnThrAsnLeuValGlu---GlyLysLeuGlnThrPro 1479
 Db 4813 CAAAGAGCTCGAGCAGTATTATTAAGTTCGAAAGCAAGAGCTGACTGCTTAGCAAGAG 4872
 QY 1480 LeuGlnAlaAspHisGlu-----GluAspSerIleAspArgArgSerGluLysMetGlu 1497
 Db 4873 TATTTGTCAGAAATGAACTGCGAACAGAGCTGACAAAGCTGCTGAGATGGAG 4932
 QY 1498 IleLysValLeuGlyLysLysLeuGluArgAsnGlnThrLeuLeuGlu----- 1513
 Db 4933 TCCAAAGTTGGCGCAGAAAGAAACAGCGAACAATCTGACTGAGCTGGAAGTACCA 4992
 QY 1514 ArgLeuGlnGluLysLysLeuGluLeuSerAsnLysLeuGluIleLeuGlnLysGluMet 1533
 Db 4993 CGACTCCAGCTACAAAGGTCTGACTTAAGTCTCGG----- 5028
 QY 1534 GluThrSerValLeuLysAspAspLeuGlnGlnLysLeuGluSerLeuLeuSerGlu 1553
 Db 5029 -----TCTTGGCTGGCATCGACACAGAACATGCTATTCAAGGC----- 5067
 QY 1554 AsnIleIleLeuLysGluAsnIleAspThrThrLeuLysHisLysSerAspThrGlnAla 1573
 Db 5068 -----CGAAATGAGAGCTGTGACATATCAAAAGAACATACTTGACAACTACAGAA 5118
 QY 1574 GlnLeuGlnLys-----ThrGlnGlnLeuGln 1583
 Db 5119 AGAACAACCAAGCATGATGTTTCATCAGATTGTGATAAAGATGCTCAGCAGACTCAT 5178
 QY 1584 Leu-----AlaLysAsnLeuAlaIleAlaAlaSerAspAsnCysProIle 1598
 Db 5179 CTAGACATGAGAAATTAAGTGAAGCTGTGCTGATTGAAACCCACAGAGAGAGTCTGCG 5238
 QY 1599 ThrGlnGluLysGluThr-----SerAla 1606
 Db 5239 GAACAGTCCCGAGATACCAATTATGAGCTTCACAGGGAAGATTAACCCAGGCTCTTCA 5298

QY 1607 AspCysValHisProLeu----- 1612
 Db 5299 GAATGCATTTTCGATGTGTCATTTTCTGTGCTCCTAATGCTTTGTCATGATTTCTCTG 5358
 QY 1613 -----GluGluLysIleLeuLeuLeuThrGluLysGlnLysGlnValThrAsnGluGln 1630
 Db 5359 GGGATATCAGAGATATTCATGATCTTCATGCTGCGGGTAAAGAGACATCAATAGAGAT 5418
 QY 1631 GlyLysLeuHisGluLysAsnGluLeuGluGlnAlaGlnValGluLeuLysGlu 1650
 Db 5419 TTGACATTACTTTCATGTATAGAGAGCCGTAGCA-----AAA 5457
 QY 1651 ValGlnHisLeuMetLysSerMetIleGluSerLysSerSerLeuGluSerLeuGlnHis 1670
 Db 5458 GTTGAAGTTCTGCTCAATGAATGAAGATTAAGACTCAAAATC----- 5502
 QY 1671 GluLysHisAspThrGluGlnGlnLeuAla-----LeuLys 1683
 Db 5503 -----CATTTACAGAGAGGTACAACTAAATGACAAATTTGAAGCATGATAGATTTGGA 5556
 QY 1684 GlnGlnMetGlnValValThrGlnGluLysGluLeuGlnGlnThrHisGlnHisLeu 1703
 Db 5557 AAAATAGTTGGGAACTTAAAGAAAGAAATCTGAGATTTAAGTGAATAATTTGGAATATTTT 5616
 QY 1704 ThrAlaGluValAspHisLeuLysGluAsnIleGluLeu-----GlyLeuAsnPheLys 1721
 Db 5617 TCTTGATACACAGAGATTTCTCAGAGAGTAAAGAACTTCTGAAAGCCCTCATTTCTGAT 5676
 QY 1722 AsnGluAlaGln---GlnLysThrThrLysGluGlnCysLeuLeuAsnGluAsnLys--- 1739
 Db 5677 TTGAAATAGCATGAGATTAATCATCAGCTGAAAGATTAATGAGATTAATGTCGSCAAAGTGTG 5736
 QY 1740 -----GluLeuGlnGlnSerGlnHisArgLeuGlnLys 1750
 Db 5737 AATGACACTGGAGAGAGATTTTCTGATGTGGAAATTAAGCTGAGTACGATCG 5796
 QY 1751 Glu-----IleGluGluLeuMetLysSerLeuLys 1760
 Db 5797 GAGAAAGCTAGCATGAGCATGAGAACCCCTGATACCTGAGAGGCGATAGAGGTATTTCA 5856
 QY 1761 AspLysGluSerAlaLeuGluThrLeuLysGluSerGluLysValIle---AsnLeu 1779
 Db 5857 ACAGAGAGCTATGTTTGAAGAAAGACAAATGAATTAAGCAGAAAGTATTTCTGCTCT 5916
 QY 1780 AsnGlnLysMetGluMetValMetLeuGluMetGluLysGluLysAsnSerGlnArgThr 1799
 Db 5917 GAGAGAGACTCTCAGTGTGCTCAGAGTGAAGAAACAGCTTCGAGAAATTTAGACT 5976
 QY 1800 ValIleAlaGluArg-----AspGlnLeuGlnAspAspLeuArgGluSerValGlu 1816
 Db 5977 ATGTCAAAATAAACACAGCGCACTGATCAGTTGCTGAAAAATGAAGAGAAACACAA 6036
 QY 1817 MetSerIleGluThr-----GlnAspAsp 1824
 Db 6037 ---GAGCTTGAAGTCTCATCAAAAGTAGTGTCTCCATTCATTCAGGTGCGACAGCAG 6093
 QY 1825 LeuArgLysAlaGlnAlaLeuGlnGlnGlnLysAspLysValGlnGluLeuThrSer 1844
 Db 6094 GTGAAGGAAAAGAGCGAACTCTTACAGCTTTGCTCTGTATGTGAGTGAAGCTGTAAAA 6153
 QY 1845 GlnIleSerValLeuGlnGluLysIleSerLeuLeuGlnGlnMetLeuTyrAsnVal 1864
 Db 6154 GACAAAACTCATCTCCAGAAAGAGCTGAGAGTTTGAAGAAAGAGCTCAGAGCACTGCT 6213
 QY 1865 AlaThrValLysGluThrLeuSerGluArgAspAsnAsnGlnSerLysGlnHisLeu 1884
 Db 6214 TTGACAAATGTGAGCTGGAACCAATTCACAACTGATTAAGAGAAAGAAATTCGCT 6273
 QY 1885 PheSerGluIleGluThrLeuSerLeuSerLeuLysGluLysGluPhe----- 1900
 Db 6274 GTCAAGGAACTGAAAGCTCAGAGCCAGACTGAGTGAATGATGATTAAGAAACCTGAT 6333
 QY 1901 -----AlaLeuGluGlnAlaGluLysAspLysAlaAspAlaIleArgLysThrIle 1917

Db 6334 GCTCCAGGCGCTTGGAGCCGCACTGCTGGAGAGAAAGTGCATTGCGATTGAGCGTGAGC 6393
 QY 1918 AspIleThrGluTyrLysSerAsnIle----- 1926
 Db 6394 TCAACACAGAGGAGAGTCACTGAGAGAGGAGGATGAGAACTGAGAGTTCGACT 6453
 QY 1927 -----GluGluGlnLeuLeuGlnGlnAlaThrAsnLeuTyrGlu 1943
 Db 6454 GAGCGCGATGAAAAGAGCAGTGCACATGCGAGAGAACTGAAAGAACCGGAGGAG 6513
 QY 1944 ArgGluSerLeuIleGlnCysGluGlnLeuAlaLeuAsnThrGluHisLeuArgGlu 1963
 Db 6514 AAGGATTCACCTT-----AAGGATAAGTTGAGAACTTGAAAGGAAATGCGAGATG 6564
 QY 1964 ThrLeuTyrSerLysAspLeuAlaLeuGlyLysMetGlnGlnAlaArgAspGluAla 1983
 Db 6565 TCAGAGAGAAAACCGAGAGTGAAGTATC-----CTTGATCCGAGAAATTCGAAAGCA-- 6615
 QY 1984 AsnLysValIleAlaLeuThrGluLysMetSerSerLeuGlnGlnLeuAsnGluAsn 2003
 Db 6616 -----GAGTGAAGACTTAAACACAAATGAGAGATG 6651
 QY 2004 ValThrThrLeuTyrGluGlu-----GlyGluTyrGlu 2016
 Db 6652 GCCAGAGCGCTGAAAGTTTGAATGATGACCTTGACGTTAAGCTCGTGAAGAAAGAAAT 6711
 QY 2017 PheTyrLeuGlnArgProSerLysGlnGlnSerSerGlnMetGluGluLeuArgGlu 2036
 Db 6712 CTGACAAAACAAATACAGAAAACAGGTCAGTTGTCAGAAACAGCAAGTTACTCTCT 6771
 QY 2037 SerLeuTyrThrLysAspLeuGlnLeuGlnAlaGluLysGluIleSerGluAlaThr 2056
 Db 6772 TCATTAAAGTCTGTAGAGAAAGAGAGCAGACAGATACGATCAAGAAAGAAATCT 6831
 QY 2057 -----AsnGluIleLysAsnLeuThrAlaLysIleSerSer 2068
 Db 6832 AAAATCGACGTGAGATGCTTCAGAAATCAGTTAAAGAGCTAAATGAGGAGTGAAGCAGC 6891
 QY 2069 Leu-----GluGluGlnIleLeuGln---AsnAlaSerIleLeuAsnGluAlaValSer 2085
 Db 6892 TTGCTGTGTGACCAAGAAATATATAGAGCCACAGACAGAGTCTAGACCCACCAATAGAG 6951
 QY 2086 GluArgLysLeuLeuArgHisSerLysGlnGlnLeuValSerGluLeu----- 2102
 Db 6952 GAAGAGCATCAGTGAAGAAATGATGAAAGCTGAGAGCCCGCTAGAAAGCTGATGAA 7011
 QY 2103 -----GlnLeuSerLeuThrLeuLysSerArgHisAlaPheAlaGlnSerLysArg 2120
 Db 7012 AAGAGCAGCTCTGTCTTACAAACACTGAG-----GAAAGTGAGCAT 7056
 QY 2121 GluLysAspGluAlaValAsnLysIleAlaSerLeuAlaGlnGluIleLysIleLeuThr 2140
 Db 7057 CATGCAAGTTACTTAAGAGGATGAGAGTGAAGAACTTGAAAGAGAGTAAATAGCCAGG 7116
 QY 2141 LysGluMetAsp-----GluPheArgAspSerLys-----GluSerLeu 2153
 Db 7117 ACAAAACCAAGAGCATGAGCTCTTGAGGACAGAAATCCAAAGAGAGAGTGAAGACCTTA 7176
 QY 2154 GlnGluGlnSerHisLeuSerGlnGlnLeuLysCysThrTyrLysThrGluLeuGlnMet 2173
 Db 7177 AAGCAAAATAGAAAGGATGACCCAAAGTCTGAGAGGCTGGAATTAAGATGTTGTTACT 7236
 QY 2174 LeuLysGlnGlnLysGluAspIleAsnAsnLysLeuAlaGluLysValLysGluValAsp 2193
 Db 7237 ATAAGGTCAGAAAAGAAATCTGACAAATGAAATTAACAAAAGAGCAAGCAGATATCT 7296
 QY 2194 GluLeuLeuGlnHisLeuSerSerLeuLysGluGlnLeuAspGlnIleGlnMetGluLeu 2213
 Db 7297 GAAATTA-----GAAATTA 7308
 QY 2214 ArgAsnGluLysLeuArgAsnTyrGluLeuCysGluLysMetAspIleMetGluLysGlu 2233

Db 7309 ATAAATTCATCATTTGAAAT----- 7329
 QY 2234 ILeSerValLeuArgLeuMetGlnAsnGluProGlnGlnGluAspValAlaGlu 2253
 Db 7330 -----ATTTCGCAAGAAAAGAGAGAGAGATGATGATGAAAGAAATCAAGCACT 7383
 QY 2254 ArgMetAspIleLeuGlnSerArgAsnGlnGlnIleGlnGlnLeuMetGluLysIleSer 2273
 Db 7384 GCCATGAGATGCTTCAACA-----CAATTAAAGAGCTCAATGACAGAGTGGCA 7434
 QY 2274 AlaValTyrSerGlnGlnHisThrLeuLeuSerSerLeuSerGlnLeuGlnLysGlu 2293
 Db 7435 GCCCTGCATAAGACCA----- 7452
 QY 2294 ThrGluAlaHisLysHisCysMetLeuAsnIleLysGluSerLeuSerThrLeuSer 2313
 Db 7453 -----GAAAGCTGTRAG-----GCCAAGAGCAGAAATCTTAGT 7485
 QY 2314 ArgSerPheGlySerLeuGlnThrGluHisValLysLeuAsnThrGlnLeuGlnThrLeu 2333
 Db 7486 AGTCAGTAGAGTGTCTGAACTTGAGAGAGGCTGCTCAAGGCTTGATGAGGCG 7545
 QY 2334 LeuAsnLysPheLysValValTyrArgThrAlaAlaValLysGluAspHisSerLeu 2353
 Db 7546 AAAAATTAATTAATTTGCAATCTCAGT-----AATGCGCTCAT 7590
 QY 2354 LysAspTyrGluLysAspLeuAlaAlaGlnGlnLysArgHisAspGlu-----Leu 2370
 Db 7591 CAAGAGTGAAGATGAGCAAGCAAGAACTGAGAGAACTGAGAAATCAAGTACAGT 7650
 QY 2371 ArgLeuGlnLeuGlnCysLeuGlnGlnHisGlyArgLysTyrSerSerPheAlaSerGlu 2390
 Db 7651 AAAAATCAAAATTCAGAACCAAGAGCAGCTGTCTTAACGTGCGAGGTGAAGAGAG 7710
 QY 2391 GluLeuLysPhe-----CysGluIleGlnPheLeuAsnGluLeuLeuPheLys 2406
 Db 7711 CACCAACTTGGAGAGAGCAAAACTAGAAATGAGAACTGACAGTGAATGAGAGAG 7770
 QY 2407 LysAlaAsnIleIleGlnSerValGlnAspAspPheSerGlnValGlnValPheLeuAsn 2426
 Db 7771 AAGATCCAGTGTACAAATCCAAAATGCTCTTTCAGAGACACATAGAAAGTGCAGG 7830
 QY 2427 GlnValGlySerThrLeuGlnGlnGlnGluLeu-----HisLysLysGlyPheMetGlnTyr 2445
 Db 7831 AGTCTTACAGAAATCTAGAGAAATGAGCTTGAATTCACAAAATGAGCAAAATGCTCT 7890
 QY 2446 LeuGlnGluPheGly-----AspLeuHisVal 2454
 Db 7891 GTTGAAAAGTAAACAAATGACTGCAGAAAGAACTGAGCTGCAGAGGAAATGATGAG 7950
 QY 2455 AspAlaLysLys-----LeuSerGlnGlyMetGlnGlnLysAsnArgGluAla 2471
 Db 7951 ATGGCACAGAAAACAGCAGAGCTGCAGAAAGAACTGAGGAGAAATATAGCTAGCT 8010
 QY 2472 SerThrIleGlnLeuLeuThrLysArgLeuLysAlaValAlaGlnSerLysIleGlnArg 2491
 Db 8011 GGAGAGTTCAGTTCTGTTGGAGAAATTAAGAGC-----AGCAAGATCA-- 8058
 QY 2492 GluIleThrValTyrLeuAsnGlnPheGluAlaLysLeuGlnGlnLysGluGlnAsn 2511
 Db 8059 -----TTGAAGAGAGCTCAACACTGAAAT 8082
 QY 2512 LysGluLeuMetArgArgMetGlnHisGlyProSerAlaSerValMetGluGlnGlu 2531
 Db 8083 AGTGAATTAAGAGAG-----AGCTAGATTCATGCACCAAGAGC 8121
 QY 2532 AsnAlaArgLeuGlnGlyIleLeuLysThrValGlnAspLysSerLysLysLeuGlnSer 2551
 Db 8122 CAGGTGAAAAGAGAGG-----AAGGTGAGAGAGAAATAGCTGAATACAGCTA 8172
 QY 2552 ArgIleLysMetLeuGlnGlnGlnLeuLeuAsnLeuValLysAspAlaMetHisLysGly 2571
 Db 8173 CGG-----CTTCATGAGAGCT 8187

QY 353 lleuaspserlysglnleugluasnleugserSerSerlythrlysalagln 372
 Db 172 AAGCAGCAAGGAGGTTTCAGTTCAGTTCAGGCTGCGCTCAAGCAAGCAAG 231
 QY 373 Alamelalysglu-----HisthrleuLeuAla 384
 Db 232 AAGGTTAAAAAAGCAAGGCTGCAAACTGAAAAAGCAAGTCAAAATGATG 291
 QY 385 GlutleuglnleuHislysgluargluasparglleThrphileuThrAsnIleval 404
 Db 292 GAATATGTGAAGTCTGCAAGAACTGACAGAAATTTCTCAT-----GAATCTCA 345
 QY 405 ValAlaserSerlynglnSerlynglnasp-----Glnarg 416
 Db 346 GTCAAGAGTCTCAAGCAATTTCCAGAAAGACACTGAATTCAGCCAAACAAATA 405
 QY 417 VallysalrglysalrglyvalThrTrpAlaProgllyslleGlnAsnSerleuHisAla 436
 Db 406 GAAAACTGGACAGCAACTTAAAGGTGAATCTGAGCTTGAAAGAACCAACAGT 465
 QY 437 SerlyValSeraspPheaspMet-----Leu 445
 Db 466 GCGCAGTCTGAGATGCTCTGCAATCCATGCAATACCAACCAAAATTTTACACT 525
 QY 446 SerArgleuProgllyAsnPheserlysalysPheSeraspMetProser----- 463
 Db 526 CCACATACCAAGTCAATATATATAGTGTCCAAAGTATGAAAGTCAAAAGAAATAT 585
 QY 464 PheProgllyleaspPheSer-----ValCysThrGluPheSer 476
 Db 586 AATAAGAGGTGAAGAAAGCAAAAGATTAGCAGCAGAGTTAAAGCTTGCAAGCTTAA 645
 QY 477 AspheaspaspAlaLeuSerMetMetAspSerAsnGlyIleaspAlaGluTrpAsnLeu 496
 Db 646 AAAGCAACCCAGACTCTCCAAAGCCACCATGATCAAGCCGACATTTGCCCGCATCAG 705
 QY 497 AlaserlysalThr-----HisArglysllyThrSerleuHis----- 509
 Db 706 GCTTATGATCTGTGTCTTCATGGCAGCAAGAAAGCCCAAGTCATCTTCACTAT 765
 QY 509 ----- 509
 Db 766 TCTCAAGAACTCAATTAGAGAGATTTCTGCACTTACTTTCTGGGAAACAGAG 825
 QY 510 -----GlnSerMetIleaspPheGlyIleSeraspSerValGlnPheHis 525
 Db 826 GTGACTCCCAAGTCGATCACTTTCGAATAGGAAAGAGATGCAATAGCAGTTCTTT 885
 QY 526 AspSerSerlysgluasnleuglnleuProlyaspSerGlyAspMetAlaGlu 545
 Db 886 GACATTTCTAGCAGTCCTCATCTTTG-----GATCAATTTAA 924
 QY 546 CysArglysalSerPheGluIleuThrSerleuGlnGlnleuGlnSerly 565
 Db 925 GGCAGAAATCAAGACTTAAGAAACAGATTAATGAGTTGGAAGCTGCGCTCAAGAGAT 984
 QY 566 GluGluGluIleuLygluIleuValGlnSerPheGluIleuLygluIleuGln 585
 Db 985 GAAAAAGAAATGAAAGCCCAAGTGAATAGTTCAA-----GAATCCAACTC 1032
 QY 586 GlnLeuSerVallysalys-----AsnleuGlnMetValThrAsnSer----- 600
 Db 1033 CAACTGAGAG-----AAAGCAAAAGTGAATTAATGAAAAAGAAAGTTTGAACAAATGT 1089
 QY 601 ArgGlnHisSerIleAsnAlaGluValGlnThrAspValGluValValArglys 620
 Db 1090 AAGGATGAATAGTGAACAAACAGCACAATAGCAG----- 1125
 QY 621 GluMetSerValleuGlyaspSerGlyTyraAsnAlaSerAsnSeraspLeuGlnaspSer 640
 Db 1126 -----CAGCGTCACCAAGATATGATGATGCAACCAAAACTGGAACAAATTTG 1173
 QY 641 SerValaspGlylysalrglyleuSerSerHisaspGluCysIleGlnHisArglyMet 660

Db 1174 ACGGAAGAT-----TTGAGTGTGCAGGACAGCAAAATGCAGAAAGTGCAGATGTTCT 1224
 QY 661 leuGlnGlnlyslleValaspLeuGlu-----GluPheIleGluAsnleuAsnlylySer 679
 Db 1225 CTGGAACAGAAATTAAGCAAAAGAAAGAGTTTCAAGGAGCTCTCCCTG----- 1278
 QY 680 GluAsnAspLygluLyserSerSerGluIleAspPheMetGluSerIleGlnleuGlu 699
 Db 1279 -----CAACACGTTCTTCCAAACACTGGACCCAGAGTGATGATCAGATG----- 1323
 QY 700 AlaIleMetAlaGluLyslasAlaAsnAlaLeuGluIleuAlaLeuMetArgAspAsnPhe 719
 Db 1324 -----AAGCCAGACTCACCCAGAGATTACAGCAAGCCAG-----AATATG 1365
 QY 720 AspAsnIleIle-----LeuGluAsn 726
 Db 1366 CACACGCTCTCCAGGCTGCACTGATTAACATCACTGATTAAGCAACAGCTAGAAAC 1425
 QY 727 -----GluThrleuysarglyleuAlaaspLeuGluArgSer----- 739
 Db 1426 AATTGGAAAGCTTTAAGCAAAAGTTGTCAGAGCTGAACAGCGTTCCAGGCGAGTCAG 1485
 QY 740 LeuLygluAsnGlnGluThrAsnGluPheGluIleleuGlnlysluThrGln----- 757
 Db 1486 ATCAAGAGAAATGACTGAGAGAACATGAGGAAATGTAAGAAAGCAAACTCCTT 1545
 QY 758 LyseGlnHisGluAlaGlnleuIleHisGlyIleGlySerleuLyseValGluAsn 777
 Db 1546 AAGAGTCACTCTGAGCAAAAGCCAGAGAAAGTCTGCCACTGAGGCAAGCTCAAGAAC 1605
 QY 778 AlaGluMet-----TyraAsnGlnAsnleuGluIleuAspLeuGluThrlystrlys 794
 Db 1606 ATCAACAGAGTTTAATATGAGCCAGAAATTTGCAAGAAATGAACCCAGAAAT--- 1662
 QY 795 LeuLeuLysluGlnGlnIleGlnleuAlaGluLeuArglysalArgAlaaspAsnleuGln 814
 Db 1663 -----ACCTCTGAGAAACATGTTAAGAGATCTTCAAGAAATAATATACAGAGAA 1716
 QY 815 LyseValArgAsnPhaspLeuSerValSerMetGlyaspSerGluIleu----- 832
 Db 1717 AACTCCTTGACTTAGAAAACTGAAGCTGTGCTGCTGATCTGGAAGAGCAGAGAT 1776
 QY 833 CysGlnGlu-----IlePheGlnleuLyserleu 843
 Db 1777 TGTTCTCAACCTTTTGAAGAAAGCAACATCACTTGAACATTTAATGATAGTTA 1836
 QY 844 SeraspAlaGluAlaValThrArgaspAlaGlnlysluGlySerPheleuArgSerGlu 863
 Db 1837 AGCAAGACAGAAAGAGTCCAAA-----GCCTTGCTGAGTCTTAGAGTTTAAAAAGAAA 1893
 QY 864 AsnleuGlnleuLygluLyserMetGluAspThrSerAsnTrpTyraAsnGlnlyslu 883
 Db 1894 GAATATGAATTAAGAAAGAGAAACTGTCTTTCTTGTGGAAGAAAGTGAACGAAAAA 1953
 QY 884 AlaIleSerleuPheGluIleuGlnleuGluThrGluLyserAsn----- 898
 Db 1954 CTTTAAT-----CAGATGATTCAGAAAGAAACTTCGAGACTAAAT 2001
 QY 899 -----TyrlysluMet 902
 Db 2002 AATCACTTGAACCTGTCTGAAGACACACAAATAAAAGTCATGATTAACACAGAGA 2061
 QY 903 GluAlaaspLeuGlnlysluGlnleuGlnSerAlaPheAsnGlnIleAsnTyraAsnGly 922
 Db 2062 GTAAGAACCTGAGAGATGACAGAGAAACCTTAAGTCTCAGATCAGAACTTCAAC 2121
 QY 923 LeuIleuAlaGlylys-----ValProArgaspLeuSerArgValGluIleuGln 939
 Db 2122 GTGTTAGACAGTATAGTCAAGTGAAGACCCAGAAACTAGCTTATATGAGACTACAG 2181
 QY 940 LyseLyseValSerGluPheSerlysluGlnleuGlnlysluAlaLeuGlnIleuLyseAsnAlaLeu 959

[illegible]

QY	1137	-----GluIleAsnAspTyrHisAsnLeuIleAlaLeuAlaThrGlu	1150
Db	3292	AAGGACGATTTCGAAGAAGAACCCAGAAATTTCTTAACAAAATTTGACTTTTGGACAA	3351
QY	1151	ArgAsn---AsnIleMetValCysLeuGluThr-----	1160
Db	3352	AGAAATCAGAAATCTGATGCTAGAGTTGGAGACAGCTCAGACAAAGCTCTGAGATCGAATG	3411
QY	1160	-----	1160
Db	3412	ACAGATAACCAAAACATTCTAAGACGACGCTGGTGTAAACGACGAAATCATGACT	3471
QY	1161	-----GluArgAsnSerLeuLysGluGlnValIleAspLeu-----AsnThr	1174
Db	3472	TTAAAGGAGAGACAAACAAATGTCGAAAAGAAAGTAATGACTTTATCAAGAGAAATGAA	3531
QY	1175	GlnLeu-----	1184
Db	3532	CAGCTGTGAAGTGAAGTAAAGACTAAACATGATCTCAAAATCTGAAATCAAGAACCAAT	3591
QY	1185	GluLysSerAspLeuGln-----LysProLysGln	1194
Db	3592	AGGAACCTCTGGAAAAGAAAAGAGAGCTGAGACAAATCAATGTAATTTTAAACCTCAGATG	3651
QY	1195	AspLeuGlnGluGlnGluValLysLeu-----LeuLeuGlnMetGlu	1208
Db	3652	GATCTGTGAAGTAAAGAAATTTCTCTGATAGATTAAATGCGCAGATGGTGCATTAAGAA	3711
QY	1209	LeuLeuLysGlnLysLeuThrAspSerGlnLeuSerIleGlnLysLeuGlnLeuGln	1228
Db	3712	GCTATTG-----CTAAGAAATTAAGGAATTTAAACTTCAGAGAAAGTGAAGAGAGAG	3762
QY	1229	LeuGluValThrGluLysLeuGlnThrLeuGlnGlnGlnMetLysAsnIleThrIleGlu	1248
Db	3763	-----GAGTGCTGCGACACATGATTAACAGACATT-----	3792
QY	1249	ArgAsnGlnLeuGlnThr---AsnProGlnAspLeuLysAlaGlnHisAspSer---Leu	1266
Db	3793	AGAGAGAGATCTTGAAMCCACCAATTTTGCAGAGACTCAGACAAAGAAATTAAGTGGCCT	3852
QY	1267	Lys-----GlnAspLeuSerGlu	1272
Db	3853	AAAGACTGTGAATATAGATGGCGAAAGAAATATATTTCAGGGCCCTCAGACTGTCAACA	3912
QY	1273	AsnIleGlnGlnSerIleGluThrGlnAspGluLeuArgAlaIleGlnGlnLeuArg	1292
Db	3913	AGTCAAAACGCAATGCACACTTTCAGTGTCTCTCCAAACCAATGACAAAGCACTGANT	3972
QY	1293	GluGlnLysGlnLeuValAspSerPheArgGlnGlnLeuLeuAspCysSerValGlyIle	1312
Db	3973	GAGCTAGGAAATATGTGTAATACACTGACAGGCTGAAAGAACTATGAACCTGTAACAGCTG	4032
QY	1313	SerSerProAsnHisAspAlaValAlaAsnGlnGluLysValSer-----	1337
Db	4033	AATGATTCAGAGCTCGAATGATATCACAGCAATCAGAAAATGCGAAGAGAGTAGGAAA	4092
QY	1338	---LeuGlnGluValAsnSerLeu---GlnSerGlnMetLeuArgGlyLysIleArgSp	1344
Db	4093	CTACTAATGAAATTAATAATATTAAATGATGACAGTGGCTTCTCCATGGTGAAGTTACTG	4152
QY	1345	Glu-----	1345
Db	4153	GAAAGCATACCAAGAGGTGAATTTGGTGAACAAACCAATGAACACACCCCTGTCTTGG	4212
QY	1345	-----	1345
Db	4213	GCTCATTTGACAGAGATTAATTCCTACAGACACTTGACATTTGTCAAGACAAAGATTCAA	4272
QY	1346	-----LeuGlnThrSerCysLysAlaLeu	1353
Db	4273	ATGCACTTTGCCAATTTGCAAGAGAAATTTCTATTCTTTAAAAAGTGAACACCAATTTTAA	4332

QY 1354 -----ValSerGluLeuLeuLeuArgLysSer 1366
 Db 4333 CATGATCAGCATCTGATGAGCTTAAATCTAGAGCTGAGACCATGTGACTCA 4392
 QY 1367 ValGluGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 1386
 Db 4393 TTAAGGCCAATAATTGCTGCTGCTCAAGATCTGAGAACTTTCAGAGCTTGCTG 4452
 QY 1387 GlysSerLeu-----GluSerGluValLeuLeuSerMetLeuGluLeuLeu 1403
 Db 4453 AAGAGATGAGCTGGCTGGAGAGAGGCTGCTCATCTCTGATCCCTCTTGTG 4512
 QY 1404 GluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 1414
 Db 4513 CCTACAGCTCTGATCTTACAGATTTGGAGACTCTCTTTACAGAGCTTTTGA 4572
 QY 1415 -----GluTyrSerSerLeuGluLeuLeuLeuLeu 1423
 Db 4573 CAGACAGAGATATCTCTTTGATTAATTTAGAAAGGCTGTTTCAAGCAAGTGC 4632
 QY 1424 SerLeuGluGluValPhe----- 1429
 Db 4633 AGTGTAGATGAAGTATTTTGCAGCATCTGACAGAGAGATCTGACCAAGAAACC 4692
 QY 1430 ---SerGlySerGluLeuValAspGluLeu-----GluValLeuLeuVal 1444
 Db 4693 CCTTCGGCCCGAGAGAGGCTGTGAAGAGCTGATCTCTGTGAGGTGTACCGGAG 4752
 QY 1445 GluLeuLeuValAlaGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 1460
 Db 4753 TCCCTCGAAGCTGAGAAAGAAATGAAAGTCAAGGATTTAGAAATTAAGAAAT 4812
 QY 1461 PheGluLeuValGluThrAlaSerThrAsnLeuValGlu---GlyLeuLeuLeuThrPro 1479
 Db 4813 CAAAGCTTCGAGCATTTTAAATTTGTAAGAGCAAGAGCTTGACTGCTTACAGACAG 4872
 QY 1480 LeuGluAlaAspHisGlu-----GluAspSerIleAspArgSerGluLeuMetGlu 1497
 Db 4873 TATTTGTCAAGAAATGAAACAGTGCACAGACAGAACTGACACGCTGCTGAGATGGAG 4932
 QY 1498 IleLeuValLeuGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 1513
 Db 4933 TCCCAAGTTGGCGGAGAAAGAAACAGACGAGCAACTGTCACTGAGCTGGAAGTACA 4992
 QY 1514 ArgLeuGluGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 1533
 Db 4993 CGACTCCAGCTCAAGAGCTGAGCTTAACTTACCTCGG----- 5028
 QY 1534 GluThrSerValLeuLeuLeuAspLeuGluLeuGluLeuLeuLeuLeuLeu 1553
 Db 5029 -----TCTTGCTGCTGATCGACACAGAAAGATGCTATTCAGGC----- 5067
 QY 1554 AsnIleIleLeuLeuGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 1573
 Db 5068 -----CGAATGAGAGCTGTGACATATCAAAAGACATCTCTGAGAACTTCACAA 5118
 QY 1574 GluLeuGluLeu-----ThrGluLeuLeuLeuLeuLeuLeuLeuLeuLeu 1583
 Db 5119 AGAAGCAAGAGATGATGCTCATCAGATTTGATTAAGATGCTCAGCAGAGCTCTCAT 5178
 QY 1584 Leu-----AlaLeuLeuLeuAlaIleAlaIleAlaSerAspSerProIle 1598
 Db 5179 CTAGACATGAGAAATACTGAGCTGTGCTGATTAACCCACAGAGAGCTCTGGG 5238
 QY 1599 ThrGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 1606
 Db 5239 GAACAGTCCCAATACATTTATGAGCTCCAGGGAGATTAACCCAGGAGCTCTCA 5298
 QY 1607 AspCysValHisProLeu----- 1612
 Db 5299 GAATGATTTCTGATTTGATTTCTGTCTTAATGCTTTGGTACCTATGATTTCTG 5358
 QY 1613 -----GluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 1630

Db 5359 GGAATCGGAGATATCATCTTCAACTCGCGGTAAAGAGCATCAATGAGAT 5418
 QY 1631 GluTyrLeuLeuHisGluTyrAsnGluLeuGluLeuAlaGluValGluLeuCysGlu 1650
 Db 5419 TTGATTAATCTCATGTGATAGAGACCGTGACAGA-----AAA 5457
 QY 1651 ValGluHisLeuMetLeuSerMetIleGluSerLeuSerSerLeuGluLeuHis 1670
 Db 5458 GTTGAAGTTGCTAATGAAATGAATTAACACTCAAACTC----- 5502
 QY 1671 GluTyrHisAspThrGluGluLeuLeuAla-----LeuLeu 1683
 Db 5503 -----CATTTACAGAGAGTCAACTATGACCAAAATGAGCATCATAGATTGAA 5556
 QY 1684 GluLeuMetGluValValThrGluLeuLeuLeuLeuLeuLeuLeuLeuLeu 1703
 Db 5557 AAAATAGTTGGGAACTTACAGAAAGAACTGAAATTAAGTAAATTAATGAAATTTT 5616
 QY 1704 ThrAlaGluValAspHisLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 1721
 Db 5617 TCTTGATACACAGAGATTAATCTCAGAGAGTAAACTCTCAAGGCTCAATCTCAT 5676
 QY 1722 AsnGluAlaGlu---GluLeuThrThrGluGluLeuLeuLeuLeuLeuLeuLeu 1739
 Db 5677 TTGAATGATCATGATTAATCATCAGTGAAGATTTGAGATATGTCGCAAGTGC 5736
 QY 1740 -----GluLeuGluLeuSerGluHisArgLeuGluLeu 1750
 Db 5737 AATGACAGCTGAGAGAGATTTCTGTATGTGAAATAGAGCTGAGATGATGATG 5796
 QY 1751 Glu-----IleGluLeuLeuMetLeuSerLeuLeu 1760
 Db 5797 GAGAAAGCTAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5856
 QY 1761 AspLeuGluSerAlaLeuGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 1779
 Db 5857 ACAGAGAGCTATGTTTAAAGAAAGCAATGAATTAAGCAAGATTAATGCTGCTT 5916
 QY 1780 AsnGluLeuMetGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 1799
 Db 5917 GAGAAAGAACTCTCAGTGTGATGATGATGATGATGATGATGATGATGATGATG 5976
 QY 1800 ValIleAlaGluLeu-----AspGluLeuLeuLeuLeuLeuLeuLeuLeu 1816
 Db 5977 ATGTCAAAAGAAAGCAAGCAGCTGATCTGATCTGATCTGATCTGATCTGATCTG 6036
 QY 1817 MetSerIleGluThr-----GluAspSer 1824
 Db 6037 ---GAGCTTGAATCTCATCAAGTGAATGATGATGATGATGATGATGATGATGATG 6093
 QY 1825 LeuArgLeuAlaGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 1844
 Db 6094 GTGAAGAAAGAAAGCAAGCTCTTCAAGCTTCTCTGATGATGATGATGATGATGATG 6153
 QY 1845 GluLeuValLeuGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 1864
 Db 6154 GACAAACTCATCTCCAGAAAGCTGACAGATTTGGAAGAGAGCTCACAGGCACTGTCT 6213
 QY 1865 AlaThrValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 1884
 Db 6214 TTGACAAATGTGAGCTGAGAAACCAATTCACAACTGATTAATTAAGAAAGATTCCT 6273
 QY 1885 PheSerGluIleGluThrLeuSerLeuSerLeuLeuLeuLeuLeuLeuLeuLeu 1900
 Db 6274 GTCAAGATCTGAAGACCTCAGGCAAGCTGATGATGATGATGATGATGATGATGAT 6333
 QY 1901 -----AlaLeuGluLeuAlaGluLeuLeuLeuLeuLeuLeuLeuLeuLeu 1917
 Db 6334 GTCTCAAGGCTTGAAGGCGGCACTGTGAGAGAAAGTGAAGTTCGATTTGAGGCTGAGC 6393
 QY 1918 AspIleThrGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 1926

Db 6394 TCACACAGAGAGAGTCATCAGCTGAGAGAGCATCGAAGAACTGAGATTGCACTT 6453
 QY 1927 -----GluGluGlnLeuLeuGlnGlnIleThrAsnLeuLysGluThrLeuTyrGlu 1943
 Db 6454 GAGCCGATGTAAGAAAGAACCTGCAATCGACAGAAAGCAAGCAAGCGGAG 6513
 QY 1944 ArgGluSerLeuLeuGlnCysLysGlnLeuAlaLeuAsnThrGlnIleLeuArgGlu 1963
 Db 6514 AATGATTCACCT-----AAGATTAAGTTGAAACCTTGAAGCAATTCAGATG 6564
 QY 1964 ThrLeuLysSerLysAspLeuAlaLeuGlnLysMetGlnGlnIleValAspGlnAla 1983
 Db 6565 TCACAGAGAAACACGAGAGCTGATGAT-----CTTGATGCCGGAATTCACAAAGCA--- 6615
 QY 1984 AsnLysValIleAlaLeuThrGlnLysMetSerSerLeuGlnGlnIleAsnGlnAsn 2003
 Db 6616 -----GAACTAGACACTTAATAACCAAAATGAAAGAGATG 6651
 QY 2004 ValThrThrLeuLysGlnGluGlu-----GlyGlnLysGlnThr 2016
 Db 6652 GCCAGAGCCCTGAAAGTTTGTGAATTAGACCTTGCACGTTAAGCTGAAAGAAAT 6711
 QY 2017 PheTyrLeuGlnArgProSerLysGlnLeuSerSerGlnMetGlnLeuArgGlu 2036
 Db 6712 CTGACAAACAAATACAGAAACAAACAGTTCAGTTCAGAACTAGACAACTTACTCT 6771
 QY 2037 SerLeuLysThrLysAspLeuGlnLeuGlnGlnIleGlnLysGlnIleSerGlnAlaThr 2056
 Db 6772 TCATTTAAAGCTCTGTAGAAAGAAAGAGCAAGCAGATACAGATTAAGAAATCT 6831
 QY 2057 -----AsnGlnLysAsnLeuThrAlaLysIleSerSer 2068
 Db 6832 AAAATGCAAGTGAAGATGCTTCAGATACATTAAGAGCCTTAATGAGCAGTACACAC 6891
 QY 2069 Leu-----GluGlnGluIleLeuGln--AsnAlaSerIleLeuAsnGlnAlaValSer 2085
 Db 6892 TTGTGTGTGTCACAAAGAAATATGACAGCCACAGAACAGCTGACCCACCAATAGAG 6951
 QY 2086 GluArgGlnAsnLeuArgGlnSerLysGlnLeuValSerGlnLeuGlu----- 2102
 Db 6952 GAAAGCATCAGCTGAGAAATAGCATTTGAAAGAGTGAAGCCGCTTAAGCTGATGAA 7011
 QY 2103 -----GlnLeuSerLeuThrLeuLysSerArgAspHisAlaPheAlaGlnSerLysArg 2120
 Db 7012 AAGAGCAGCTGCTGTCTTCAACAACTGAAG-----GAAAGTACGAT 7056
 QY 2121 GluLysAspGlnAlaValAsnLysIleAlaSerLeuAlaGlnIleLysIleLeuThr 2140
 Db 7057 CATGCAGATTACTTAAGGTAGAGTGAGAACTTGAAGAGAGCTGAGATAGCCAGG 7116
 QY 2141 LysGlnMetAsp-----GluPheArgAspSerLys-----GluSerLeu 2153
 Db 7117 ACAACCAACAGACATGCAAGCTCTTGAGCAGAGAAATTCACAAAGAGAGCTGAGACCTTA 7176
 QY 2154 GlnGlnGlnSerSerHisLeuSerSerGlnGlnLeuGlnCysThrTyrThrGlnLeuGlnMet 2173
 Db 7177 AAACCAAAATAGAGAGGATGACCCAAAGCTGAGAGCTGGAATTAAGATTGTTACT 7236
 QY 2174 LeuLysGlnGlnLysGlnAspIleAsnAsnLysLeuAlaGlnLysValLysGlnValAsp 2193
 Db 7237 ATAAGGTGACAGAAAGAAATCTGACAAATGAATTACAAAGAGCAAGCAGATATCT 7296
 QY 2194 GluLeuLeuGlnHisLeuSerSerLeuLysGlnGlnLeuAspGlnIleGlnMetGlnLeu 2213
 Db 7297 GAATTA-----GAAATA 7308
 QY 2214 ArgAsnGlnLysLeuArgAsnTyrGlnLeuGlnCysGlnLysMetAspIleMetGlnLysGln 2233
 Db 7309 ATAATTCATCATTTGAAAT----- 7329
 QY 2234 IleSerValIleLeuArgLeuMetGlnAsnGlnProGlnGlnGlnLysAspAspValAlaGlu 2253
 Db 7330 -----ATTTTGCAAGAAAGAGCAAGAGAAAGTACAGATGAAGAAATCAAGCACT 7383

QY 2254 ArgMetAspIleLeuGlnSerArgAsnGlnGlnIleGlnGlnLeuMetGlnLysIleSer 2273
 Db 7384 GCCATGAGATGCTTCAACA-----CAATTAAGAGCTCAATAGAGAGTGGCA 7434
 QY 2274 AlaValTyrSerGlnGlnHisThrLeuLeuSerSerLeuSerSerGlnLeuGlnLysGlu 2293
 Db 7435 GCCCTGCATTAATGACCA----- 7452
 QY 2294 ThrGlnAlaHisLysHisCysMetLeuAsnIleLysGlnSerLeuSerSerThrLeuSer 2313
 Db 7453 ---GAAGCCTGAAAG-----GCCAAAGACAGAAATCTTGTAGT 7485
 QY 2314 ArgSerPheGlySerLeuGlnThrGlnHisValLysLeuAsnThrGlnLeuGlnThrLeu 2333
 Db 7486 AGTCAAGTAGAGTCTCTGAACCTTAGAAGGCTCAGTTGCTTAACAAGCCTTGATAGGCC 7545
 QY 2334 LeuAsnLysPheLysValValTyrArgThrAlaAlaValLysGlnAspHisSerLeuIle 2353
 Db 7546 AAAATTAATTAATGTTTGTGCAATCTCAGTG-----AATGGCCTCAAT 7590
 QY 2354 LysAspTyrGlnLysAspLeuAlaAlaGlnGlnLysArgHisAspGlu-----Leu 2370
 Db 7591 CAAGAACTAGAAAGATGCCAAGCAGAACTGGAAGAAAGATGAAGAAATCAGTACACTG 7650
 QY 2371 ArgLeuGlnLeuGlnCysLeuGlnGlnHisGlyArgLysTyrPheAspSerAlaSerGlu 2390
 Db 7651 AAAATCAAAATTCAGAACCCAGACAGCAGCTTGTCTTAACGCTCCAGCTGGAAGAGAG 7710
 QY 2391 GluLeuLysPhe-----CysGlnIleGlnPheLeuAsnGlnLeuLeuPheLys 2406
 Db 7711 CACCAACCTTGGAAAGAGCAAAATCTTAGACATGAGAAATTCAGACGTGAATGGAGCAG 7770
 QY 2407 LysAlaAsnIleIleGlnSerValGlnAspPheSerGlnValGlnValPheLeuAsn 2426
 Db 7771 AAGATCAAGTGCCTCAATCCAAATGCTCTTGCAGACACATTAGAACTGCTGCAG 7830
 QY 2427 GlnValGlySerThrLeuGlnGlnLeuGlnGlu--HisLysLysGlyPheMetGlnThr 2445
 Db 7831 AGTTTTCACAAAGATCTAGAGAAATGAGCTTCAATGACAAAAATGACAAATGCTCTT 7890
 QY 2446 LeuGlnGlnPheGly-----AspLeuHisVal 2454
 Db 7891 GTTGAAGAAATGAACAAATGACTCAAGAGAACTGAGTCAGAGGAAATGATGATGAG 7950
 QY 2455 AspAlaLysLys-----LeuSerGlnLysMetGlnGlnLysAsnArgArgIleAla 2471
 Db 7951 ATGCCACAGAAACAGACAGAGCTGCMAAGAACTCACTGAGAGAAATTAAGCTTACT 8010
 QY 2472 SerThrIleGlnLeuLeuThrLysArgLeuLysAlaValAlaGlnSerLysIleGlnArg 2491
 Db 8011 GGAGAGTTGCACTTACTTGTGGAAGAAATTAAGAC-----ACGAAAGATCA--- 8058
 QY 2492 GluIleThrValTyrLeuAsnGlnPheGlnAlaLysLeuGlnGlnLysGlnGlnAsn 2511
 Db 8059 -----TTGAAGAGACTCACACTAGAAAT 8082
 QY 2512 LysGlnLeuMetArgArgMetGlnHisHisGlyProSerAlaSerValIleMetGlnGln 2531
 Db 8083 AGTGATTAAGAGAG--ACCTTAGATTCATGATCACAAGAAC 8121
 QY 2532 AsnAlaArgLeuLeuGlyTyrLeuLysThrValGlnAspGlnSerLysLeuGlnSer 2551
 Db 8122 CAGGTGAGAAAGAGAGG-----AAAGTGAAGAGAAATAGCTGAATATCATGACTA 8172
 QY 2552 ArgIleLysMetLeuGlnAsnGlnLeuAsnLeuValLysAspAspAlaMetHisLysGly 2571
 Db 8173 CGG-----CTTCATGAACT 8187
 QY 2572 GluLys-----ValAlaIleLeuGlnAspLysLeuLeuSerArgAsnAlaGlnAlaGlu 2589
 Db 8188 GAAAGAAACACCAAGCCTTGTCTTTGGAC-----ACAAACCAAGCTATGAAAGTGAAG 8241

```

QY 2590 LeuAsnAlaMetGlnValLysLeuThrLysGlnAspAsnLeuGlnAlaAlaMetLys 2609
Db 8242 ATCCAGACATACCGAGAGAAATGACTTCTAAGAGAAATGCTCAGTCCACAGAGCTG 8301
QY 2610 GlnIleuLysLeuGlnLysMetValAlaLysGlyValProLysGlnGluIle 2629
Db 8302 GAGATTAAGACCTTTAAAGTCTACT-----AAAGAGAGCTC 8337
QY 2630 AspAsn-----LeuLysThrLysValLysIleGluMetGlu 2642
Db 8338 AATTAATTCATTAAGAGCTACTACTAGATTTTGGAGAAATTAAGAAACCAAGATGAGC 8397
QY 2643 LysIleLysIle-----SerLysAlaThrAspGlnGluIle 2654
Db 8398 AATCTAAATATGTAAATCAAGTTGAGAGAGAAATGACGTCGCCGAGGAGAAATGAG 8457
QY 2655 AlaThrLeuLysSerCys-----LeuGlnAspLysGlnGluIleValAspArgLys 2672
Db 8458 TTGTTGATCAAAATCTCTTAACACAGCTGAGAGAGAAAGAGATCTGCAGAAAGACTC 8517
QY 2673 GlnIleuLysArgAlaGlnAlaAspAsnAspThrValCysValProLysAspPyr 2692
Db 8518 TCTCAACTTCAAGCTGCACAGAGAGAGAGAGAAACA----- 8553
QY 2693 GlnLysAlaSerThrPheProValThrCysGlyGlySerGlyIleValGlnSerThr 2712
Db 8554 -----GGTACTGTTATGATACC 8571
QY 2713 AlaMetLeuValLeuGlnSerGlu-----LysAlaAlaLeuGlu----- 2725
Db 8572 AAGGTCGATGATTAACAACAGATCAAGACTGAAAGAAAGAACTGTAAGAAAGAAACC 8631
QY 2726 ArgGlnLeuSerHisThrLysLysLysThrHisLysLeuSerArgThrMetSerSer 2745
Db 8632 AAGGAGGACATGAATTAATCTGATTAAGTACTGCTCTTG-----CTATAAGCAT 8682
QY 2746 GlnAspArgLysLysThrLysAlaLysSerAspAlaHisSerSerHisThrGlySer 2765
Db 8683 GAAAGTTAGAGAAAGCTAAAGAGATGTTAGACACAGAGGCTGCTGTCACAG 8742
QY 2766 His-----ArgGlySerProHisLysThrGluThrValArgHisGlyPro 2780
Db 8743 CAATCTAAACAAGATTCGAGAGGCTCTCT-----TTCCTAGGTCA 8784
QY 2781 ValThrProGlnArgSerGluMetProSerLeu-----HisLeuGlySerProLys 2797
Db 8785 GTTGTTCAGACATCTCCATCTCTCTGTTACTGAAAGAGTGTATCATCTGGCCAA 8844
QY 2798 LysSerGlnSerSerThrLysArgValLys----- 2808
Db 8845 AATTAAGCTTCAGAGAGAGAGAGAGATCCAGTGAATATGGAGAAATGGTGAGAGACCA 8904
QY 2809 -----ProAsnArgSerGluIleLysSer-----GlnLeuValMetSer----- 2821
Db 8905 ACACTGCTACCCACAGAGAGCTTTCTATAAAAGCAGAAAGACATGATGAGTGTAT 8964
QY 2822 -----ProGlyLys----- 2824
Db 8965 CACCTCGAGAGACAGAGAGAGTACTGAGTTCAGCCAGAGAGACTTCACAGACTTGA 9024
QY 2825 -----ThrGlyMetHisLys-HisIle----- 2831
Db 9025 AAGAAAGGGTTGGCTGACATCCGACAGAGAAAGACTAGCCCATATATCTCTCGAAAGACA 9084
QY 2832 -----LeuSerProSerLysValGlyLeuHisLysLysAlaGlnAlaLeuSerPro-- 2847
Db 9085 ACCATGGCACTCGACACAGCCCCCTGGCTGACAG-AGGTAAAGCCTATCCCAT 9143
QY 2848 -----AsnArgSe 2850
Db 9144 GAGTCTCGGAAAGAAATCTTCAGAGTCTCAACACCAAGCTGTGCGAGAGAGATC 9203
QY 2850 rgluMetProThrGlnHisValLysSerProGlyLysThrGly-----LeuHisLysAs 2868

```

```

Db 9204 ACAAAAGCTCAAGTGTGCTCAGCGAGGCCAGTAGATTCAGGACCATCTCCAGAGAAC 9263
QY 2868 nLeuThrGlnSerThrLysPheAspAsnLeuSerSerProCysLysGlnGlnValGln 2888
Db 9264 CACCACGAATTCGTCGCCAGTCAATATCTCTGAGAGAAAGTCCGACTGCAGCCCCAG 9323
QY 2888 nGlnAsnLeuAsnSerProLysGlyLysLeuPhe-----AspValLysSe 2903
Db 9324 AAGAGGCTCGAGGCTCAGCGAGGCCGACTTGTCCAGCCCAAGCTGCAGCTGAGATC 9383
QY 2903 rLysSerMetProLysCysProSerGln 2912
Db 9384 CAAGGCGACTGAGAACTGTAAGTCCAG 9411

RESULT 7
US-09-967-768A-186
/ Sequence 186, Application US/09967768A
/ Patent No. US20020150877A1
/ GENERAL INFORMATION:
/ APPLICANT: Augustus, Meena
/ TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sigs
/ FILE REFERENCE: 689290-72
/ CURRENT APPLICATION NUMBER: US/09/967,768A
/ PRIOR FILING DATE: 2001-09-28
/ PRIOR APPLICATION NUMBER: US/60/236,109
/ PRIOR FILING DATE: 2000-09-28
/ PRIOR APPLICATION NUMBER: US/60/236,034
/ PRIOR FILING DATE: 2000-09-28
/ PRIOR APPLICATION NUMBER: US/60/236,111
/ NUMBER OF SEQ ID NOS: 325
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 186
/ LENGTH: 10211
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-967-768A-186

Alignment Scores:
Pred. No.: 7,99e-64 Length: 10211
Score: 1139.00 Matches: 703
Percent Similarity: 37.36% Conservative: 556
Best Local Similarity: 20.86% Mismatches: 1052
Query Match: 7,71% Indels: 1060
DB: 10 Gaps: 139

US-09-150-867-1 (1-2954) x US-09-967-768A-186 (1-10211)
QY 333 ThrProHisValAsnGlnValLeuAspAspGlnAlaLeuLysArgLysGln 352
Db 112 ACAAGAGCTCTTCAGAAATTCAGAGCTTGAAGAGACGTTGACAACTGAGAAAGGAA 171
QY 353 IleLeuAspLeuLysGlnLeuGlnAsnLeuGlnSerSerGluThrLysAlaGln 372
Db 172 AAGCAGCAAGAGCGAGTTCAGCTTGACAGTTCGAGGCTCGCCTCAGAGCAAAACAG 231
QY 373 AlaMetAlaLysGlnGlu-----HisThrGlnLeuAla 384
Db 232 AAGCTTGAATTAAGAAACCGAGGTAACAACCTGAAAGAGGAGAACTCAAAATTCATG 291
QY 385 GlnIleLysGlnLeuHisLysGlnArgGlnAspArgLysLeuThrAsnIleVal 404
Db 292 GAATTAATGTCAAACTCGAGAGAAACCTAAGCAGAGATTTCTCAT-----GAACTTCAA 345
QY 405 ValAlaSerSerGlnGlnSerGlnAsp-----GlnArg 416
Db 346 CTCAGAGAGTCAACAAGTGAATTCAGAGAGAGCACTGAATTCAGGCAAAACCAATA 405
QY 417 ValLysArgLysArgValThrThrAlaProLysLysIleGlnAsnSerLeuHisAla 436
Db 406 GAAAAACTGGAACAGAACTTAAAGGTGAATCTAGCTTGAAGAGGCAACAAAGCT 465

```

```

Oy 437 SerGlyValSerAspPheAspMet-----Leu 445
    : : : : :
Db 466 GCGCAGTGTGAGATGTCTCTGTGATCCATGCAATACACCAAAAAATTTTTCACAACT 525
Oy 446 SerArgLeuProGlyAsnPheSerLysAlaLysPheSerAspMetProSer-----463
    : : : : :
Db 526 CCACATACCAAGTCATATTATATAGTGTCCAGTATGAAAGCTTAAGAAAAAATAT 585
Oy 464 PheProGluIleAspAspSer-----ValCysThrGluPheSer 476
    : : : : :
Db 586 AATAAAGGCTTGAAGAAAGAAAAAGATTAGAGCAGAGGTTAAAGCCTTGAGCAGCTAA 645
Oy 477 AspPheAspAlaLeuSerMetMetAspSerAsnGlyIleAspAlaGluThrPasnLeu 496
Db 646 AAAGCAAGCCAGACTCTTCCACAGCCACCATGAATCACCGGACATTTGCCGACATCAG 705
Oy 497 AlaSerLysValThr-----HisArgGluLysThrSerLeuHis-----509
    : : : : :
Db 706 GCTTCATCATCTGTGTCTCATGCGACAGAGAAAGACCCCAAGTCATCTTTCATCTAAT 765
Oy 509 -----509
Db 766 TCTCAAGAACCTCAATTAGAGAGATTCTCTGCATCTTACTTTTCTGGGACACAGAG 825
Oy 510 -----GlnSerMetIleAspPheGlyGlnIleSerAspSerValGlnPheHis 525
    : : : : :
Db 826 GTGACTCCAAAGTCCATCTTGCATTAAGGAAAAGAGATGCTTAATAGCAGTTTCTTT 885
Oy 526 AspSerSerLysGluAsnGlnLeuGlnLysLeuProLysAspSerGlyAspMetAlaGlu 545
    : : : : :
Db 886 GACATCTCTACGAGCTCTCATCTTTTG-----GATCAATATAAA 924
Oy 546 CysArgLysAlaSerPheGluLysGluLysIleThrSerLeuGlnGlnLeuGlnLys 565
    : : : : :
Db 925 GCGCAAGAAATCAAGAGCTAAGAAACAAAGATTAATGAGTTGGAAGCTGACGCTGCAAGACAT 984
Oy 566 GlnGluGlnLysLysGluLeuValGlnSerPheGluLeuLysIleAlaGluLeuGlnLys 585
    : : : : :
Db 985 GAAAAAGAAATGAAAGCCAGAGTGAATAAGTTTCA-----GAACTCCAACTC 1032
Oy 586 GlnLeuSerValLysAlaLys-----AsnLeuGlnMetValThrAsnSer---600
    : : : : :
Db 1033 CAATCTGGAG-----AAAGCAAAAGTGAATTAATTGAAAAAGAAATTTTGAACAAATGT 1089
Oy 601 ArgGlnHisSerIleAsnAlaGluValGlnThrAspValGluLysGluValAlaArgLys 620
    : : : : :
Db 1090 AGGATGAACTAGTGAGAACAAACAGACAAATACGAC-----1125
Oy 621 GluMetSerValLeuGlyAspSerGlyTyrAsnAlaSerAsnSerAspLeuGlnAspSer 640
    : : : : :
Db 1126 -----CAGGCCTCAACCAAGTATACTGCATTTGGAACAAAAAAGTAAAAAATG 1173
Oy 641 SerValAspGlyLysArgLeuSerSerSerHisAspGlyCysIleGlnHisArgLysMet 660
    : : : : :
Db 1174 ACGGCAAGAT-----TTGAGTTGTGACAGACAAATAATGCAAGAAAGTGGCAGATGTTCT 1224
Oy 661 LeuGlnGlnLysIleValAspLeuGln-----GluPheIleGluAsnLeuAsnLysLysSer 679
    : : : : :
Db 1225 CTGGAACAGAAATTAAGGAAAAAGAAAGAGTTTCAAGAGAGAGCTCTCCGT-----1278
Oy 680 GluAsnAspLysGlnLysSerSerGlnGlnAspPheMetGlySerIleGlnLeuGlnGlu 699
    : : : : :
Db 1279 -----CAACAGCGTCTTCTTCCAAACACTGACCAAGAGATGATCCAGATG-----1323
Oy 700 AlaIleMetAlaGluLysAlaAsnAlaLeuGlnGluLeuAlaLeuMetArgAspAsnPhe 719
    : : : : :
Db 1324 -----AAGGCCAGACTCACCCAGAGTTACAGCAAGCAAG-----AATATG 1365
Oy 720 AspnIleIleLys-----LeuGlnAsn 726
    : : : : :
Db 1366 CACAACGTCCTGACAGCTGAACTGATTAACCTACATCATGTAAGCAACAGCTAGAAAAAC 1425

```

```

Oy 727 -----GluThrLeuLysArgGluIleAlaAspLeuGluArgSer-----739
    : : : : :
Db 1426 AATTGGAAAGACTTTRAGCAAAAGTTGTGCACAGCTGAAACAGGCTTCCAGCCAGTCAAG 1485
Oy 740 LeuLysGluAsnGlnGluThrAsnGluPheGluIleLeuGlnLysGluThrGln-----757
    : : : : :
Db 1486 ATCAAGAGAAATAGACTGAGAGAGAAAGATGAGAAATGAAAGAAACAAACCTCTT 1545
Oy 758 LysGluHisGluAlaGlnLeuIleHisGluIleGlySerLeuLysLeuValGluAsn 777
    : : : : :
Db 1546 AAGAGTCACTGTGACCAAAAGGCCAGAGAAAGTCTGCCACTGTGAGCAGACACTCAAGAAC 1605
Oy 778 AlaGluMet-----TyrAsnGlnAsnLeuGlnGluAspLeuGluThrLysThrLys 794
    : : : : :
Db 1606 ATCAACAGACTGTTTAATCAGAGCCAGAAATTTTGCAGAAAGAAATGAAACCGAAGAT---1662
Oy 795 LeuLeuLysGluGlnGluIleGlnLeuAlaGluLeuArgLysArgAlaAspAsnLeuGln 814
    : : : : :
Db 1663 -----ACCTTCAGAGAAACCATGTTAAGAGATCTTCAAGAAATAATATCAGCAGAA 1716
Oy 815 LysLysValArgAsnPheAspLeuSerValSerMetGlyAspSerGluLysLeu-----832
    : : : : :
Db 1717 AACTCCTTGACTTTAGAAAAAACTGAAGCTTGCTGTGCTGATCTGAAAGACGACGAGAT 1776
Oy 833 CysGluGlu-----IlePheGlnLeuLysGlnSerLeu 843
    : : : : :
Db 1777 TGTTCCTCAAGCCTTTTGAAGAAAAAGACAACTCATTCATGAACACTTAATGATTAAGTTA 1836
Oy 844 SerAspAlaGluAlaValThrArgAspAlaGlnLysGluCysSerPheLeuArgSerGlu 863
    : : : : :
Db 1837 AGCAAGACAGAGAAAGAGATCCAAA---GCCCTGCTGAGTCTTGAAGTTAAAAAGAAA 1893
Oy 864 AsnLeuGluLeuLysGlnLysMetGluAspThrSerAsnTyrPyrAsnGlnLysGluLys 883
    : : : : :
Db 1894 GAATATGAAATTAAGAAAGAGAAAACTCTGTTTCTTGTGAAAGTGAAGAAAAAGAAAAA 1953
Oy 884 AlaAlaSerLeuPheGluLysGlnLeuGlnLysThrGlnLysSerAsn-----898
    : : : : :
Db 1954 CTTTAACT-----CAGATGAAATCGAAAAAGAAACCTTCAGACGATTAAT 2001
Oy 899 -----TyrLysLysMet 902
    : : : : :
Db 2002 AATCACTTGAACCTGTCTGAAAGACACAGCAAAATTAAGATCAATATACAAACAGAGA 2061
Oy 903 GluAlaAspLeuGlnLysGlnLeuGlnSerAlaPheAsnGluIleAsnTyrLeuAsnGly 922
    : : : : :
Db 2062 GTAAGACGCTGGAGATGACAGAGAAAACTTAAGTGTGAGATCAGAAACCTTCAACAA 2121
Oy 923 LeuLeuAlaGlyLys-----ValProArgAspLeuLeuSerArgValGluLeuGln 939
    : : : : :
Db 2122 GTGTTAGACAGTAAGTCAAGTGAAGTGAAGACCCAGAAACCTTAATATGAGAGCTACAG 2181
Oy 940 LysLysValSerGluPheSerLysGlnLeuGlnLysAlaLeuGlnGluLysAsnAlaLeu 959
    : : : : :
Db 2182 CAGAAA---GCTGAGTGTCA-----GATCAGAAACATCAGAAAG 2217
Oy 960 GluAsnGluValThrCysLeuSerGluTyrLysPheLeuProAsnGluValGluLysLeu 979
    : : : : :
Db 2218 GAAATAGAAATATATGTGTTGAAGACTTCTCAG---CTTACTGGCGAAGTTGAAGATCTA 2274
Oy 980 LysAsnGlnIleSerLysAlaSerGluGluIleMet-----991
    : : : : :
Db 2275 GAACACAGCTTCACTTACTGTCAATATGAATTAATGAAACAAAGACCGGTTACCAAGAC 2334
Oy 992 -----LeuLeuLysGlnGluLys-----997
    : : : : :
Db 2335 TTGATGCCGATATATAGAGCCTCAGGATGCTGTAATAATCCAAAGATGCTTCTGTGTG 2394
Oy 998 -----GluHis 999
    : : : : :
Db 2395 ACAATGAAATCATCATCAGAGAAAGTCTTTGCTTTTGTATCAGACCGTCCATGATCAT 2454
Oy 1000 Ser---AlaSerIleIleSerLysGlnGluIleIle-----1010

```

Db 2455 TCCTTGCATAATTAATGAGAACAGAGCATGCTTCAGAGAGAGTGAATGCTCGT 2514
 Qy 1010 -----LysProlysglu 1194
 Db 2515 TTGAAGACAGACCAAAATGCGAAAAATTCGCATCTCTCAAAAATAGAGTTGATCACTT 2574
 Qy 1011 -----MetGlnGlnGlnSerGlnGlu 1017
 Db 2575 GAATTTTCATTAGAGTCTCAAAAAACAGATGACACTCAGACCTCGCAAAAGCGTGCGAAGAG 2634
 Qy 1018 ILLeuGlnLeuThrAspGluValThrHisThrGlnSerLysValGlnGlnThrGlnGlu 1037
 Db 2635 TTGGTCGAATTCAGAGAGAAATAGAGAAATCTCATGAAACAGACAGATGATCA 2694
 Qy 1038 GlnThrLeu-----GlnMetLysMetHisAsp 1048
 Db 2695 AGTTTGTGCTGAAACAGACAGCCGATTAAGTTACAGAGAGACACTTCTGCTCAC 2754
 Qy 1048 ----- 1048
 Db 2755 CAGAACTGTGCTGAACCTTAAGTCCCTTGAGACAGAGAAAAAGAGCTGCAACTT 2814
 Qy 1049 LeuPheGlnLysTyrIleArgAsnLysSerGlu----- 1059
 Db 2815 TTAAATGATTAAGTAGAAGAACTGAGCAGCAGAGATTCAGAAATTAAGAGCAACCAT 2874
 Qy 1060 ---AlaGlnAspLeuLeuArgGlnMetGlnAsnLeuLysGlnThrMetGlnSerValGlu 1078
 Db 2875 CTACTTGAAGCTCTTAAGAGAGCTTAACTTAACTCGAAACCCCA---AGCTTGAG 2931
 Qy 1079 ValLys-----IleAlaAspThrLysHisGlnLeuGlnGlnThrIleArg 1093
 Db 2932 AAGAAAGAAATGAGATTCATCATCTTCTTAATTAAGGAAATTAAGAGAGCTGACCCA 2991
 Qy 1094 Asp-----LysGlu-----GlnLeuLeuHisGlnLysLysTyrPhePhe 1106
 Db 2992 GAGATGAGGACTCTTAAGAAATTAATGATTCATTAATCAAGAGAGAGAACTTAATC 3051
 Qy 1107 GlnAlaMetGlnThrIlePhe-----ProIleThrProLeu 1118
 Db 3052 CAGAAAGAGTGAAGCTTTGCAAACTATATGATGAAAGGAGAAAGCAATTCAGAGTAA 3111
 Qy 1119 SerAspSerLeuProProSerLysLeuVal-----GlnGlyAsn 1131
 Db 3112 TCTGTGATGATCAAGCAAGAAATTAATTTACTACAAAGATGTAAGAAACCGGAAT 3171
 Qy 1132 SerGlnAspProIle----- 1136
 Db 3172 GCATATGAGATCTTAAGTCAAAATACAAAGCAGACAGAAAGAAATTTAAATTGAA 3231
 Qy 1136 ----- 1136
 Db 3232 TGCCTTGCATAATGATGCACTAGTCTTTGTAATAATAGAAATAGATTGGACAGCTA 3291
 Qy 1137 -----GlnIleAsnAspTyrHisAsnLeuIleAlaLeuAlaThrGlu 1150
 Db 3292 AAGGAAGCATTTGCAAGAGAACACCAAGAAATTTTACAAAAATTAAGCAATTTGCGAAGAA 3351
 Qy 1151 ArgAsn---AsnIleMetValCysLeuGlnThr----- 1160
 Db 3352 AGAAATCAGAAATCTGATCTAGAGTGGAGACAGTGCAGCAAGCTCTGAGATCTGAGATG 3411
 Qy 1160 ----- 1160
 Db 3412 ACAGATACCAAAACATTTCTAAGAGCGAGCTGGTGTAAAGCAAGAAATCATGACT 3471
 Qy 1161 -----GluArgAsnSerLeuLysGlnValIleAspLeu-----AsnThr 1174
 Db 3472 TTAAGAGAGAAACAAACAAATTAATGACTTAATTAAGAGAGAAATGAA 3531
 Qy 1175 GlnLeu-----GlnSerLeuGlnAlaGlnSerIle 1184

Db 3532 CAGCTGATGAAGATTAATGAAGACTAAACATGAAATGTCAAAATCTAGAAATCAGAACCAATT 3591
 Qy 1185 GlnLysSerAspLeuGln-----LysProlysglu 1194
 Db 3592 AGGAACCTCTGTAAAGAAAGAGAGAGAGAGAGAAATCAATGTAATTTAAACCTAGATG 3651
 Qy 1195 AspLeuGlnGlnGlnValValLysLeu-----LeuLeuGlnMetGlu 1208
 Db 3652 GATCTTGAAAGTAAAGAAATTTCTAGATAGTTAAATGCGCAGTGTGTGCAATTAAGA 3711
 Qy 1209 LeuLeuLysGlnHisLeuThrAspSerGlnLeuSerIleGlnLysLeuGlnLeuLysn 1228
 Db 3712 GCTATG-----CTAAGAAATTAAGGATTAACCTTCAGAGAAAGTGAAGAGAGAG 3762
 Qy 1229 LeuGlnValThrGlnLysLeuGlnThrLeuGlnGlnGlnMetLysAsnIleThrIleGlu 1248
 Db 3763 -----GAGTGCCCGCAGCATGAAATTCAGACAAATT----- 3792
 Qy 1249 ArgAsnGlnLeuGlnThr---AsnPheGlnAspLeuLysAlaGlnHisAspSer---Leu 1266
 Db 3793 AGAGAGATCTTGAAACAGACCAATTTGCAAGACATGCACTCAGAGAAATTAATGAGCCCTT 3852
 Qy 1267 Lys-----GlnAspLeuSerGlu 1272
 Db 3853 AAAGACTGGAATTAATGATCGGAGAGAAAGTATTTTCAGGGCCCTCATGAGTTGCAACA 3912
 Qy 1273 AsnIleGlnGlnSerIleGlnThrGlnAspGlnLeuArgAlaGlnGlnGlnLeuArg 1292
 Db 3913 AGTCAAAGACAGCAATGACACACTTCACTGCTCTCTCAAAACAAATGAAAGAGTGAAT 3972
 Qy 1293 GlnGlnLysGlnLeuValAspSerPheArgGlnGlnLeuLeuAspLysSerValGlyIle 1312
 Db 3973 GACCTGAGAGAAATTAATGTAATGATGACAGGCTGAAAGCTAATGACTGTAACCTGACCTG 4032
 Qy 1313 SerSerProAsnHisAspAlaValAlaAsnGlnGlnLysValSer----- 1327
 Db 4033 AATGATTCAGAGTCAAGATATATACACAGCACTAGAGAAATAGCAGAAAGAGAGGAA 4092
 Qy 1328 ---LeuGlyGlnValAsnSerLeu-----GlnSerGlnMetLeuArgGlnGlnValAsp 1344
 Db 4093 CTACTTAATGAAATTAATTAATTAATTAATGATGACAGTGGCTTCTCATGAGTGAATGAGT 4152
 Qy 1345 Gln----- 1345
 Db 4153 GAAGCATATACAGAGAGTGAATTTGGTGAACAACAAATGACAGACACCCTGTGCTTGG 4212
 Qy 1345 ----- 1345
 Db 4213 GCTCCATTGAGAGAGATTAATCTCTACAGACACTGACATTTGCAGACAAAGATTCAA 4272
 Qy 1346 -----LeuGlnThrSerCysLysAlaLeu 1353
 Db 4273 ATGCATTTGCCGAAATTCAGAGAAATTTCTTAATCAAAAGTGAACACAAATTTTA 4332
 Qy 1354 -----ValSerGlnLeuGlnLeuLeuArgAlaHisValLysSer 1366
 Db 4333 CATGATAGACACTGTCAGATGAGACTCTTAATATGTCAGAGCTGAGACCTATGTTGACTCA 4392
 Qy 1367 ValGlnGlnGlnLeuGlnValIleThrLysLysLeuAsnGlnLysGlnGlnIleLeu 1386
 Db 4393 TTAAGGCCGAAATTTGCTTGTCTCAACAGAACTGGAACACTTTCAAAGTGAAGTGGTG 4452
 Qy 1387 GlyLysSerGlu-----GlnSerGlnValLeuLysSerMetLeuGlnAsnLeuLys 1403
 Db 4453 AAGGAGATGAGCTGGCTTGAGAGAGAGGAGGCTGCTTCATCTCCTGATCTCTTGTGTG 4512
 Qy 1404 GlnAspAsnAsnLysLeuLysGlnGlnAlaGlu----- 1414
 Db 4513 CCTGACAGCTCTACTTACAGAGTTTGGAGAGACTCCCTTTTACAGAGCTCTTTAGAA 4572
 Qy 1415 -----GlnTyrSerSerLysGlnAsnGlnPhe 1423
 Db 4573 CAGACAGAGATATGCTCTTTGAGATTAATTTGAAGAGGCTGTTTTCAGCAACCACTGC 4632

QY 1424 SerLeuGluValPhe----- 1429
 DB 4633 AGTGTAGTGAATTTTGGACAGCTGCAGAGAGAAATTCGACGAGAAACAC 4692
 QY 1430 ---SerGlySerGlnLysLeuValAspGluLeu-----GluValLeuYsAla 1444
 DB 4693 CCTTCGGCCCCCAGGAGGGGTGTAAGACCTTGAGTCCCTCTGAGAGGTGACCGGAG 4752
 QY 1445 GlnLeuYsAlaAlaGluGluArgLeuGlu-----LysAspArgPyr 1460
 DB 4753 TCCCTCGAGAGCTAGAGAGAAATGCAAGATTAATGAAAATAAGAAATTT 4812
 QY 1461 PheGluLeuValGlnThrAlaAsnThrAsnLeuValGlu---GlyLysLeuGluThrPro 1479
 DB 4813 CAAGAGCTCGACAGATTATTAAGTTCTGAAGCAAGAGCTTGACTGCTTGAAGAGCAG 4872
 QY 1480 LeuGlnAlaAspHisGlu-----GluAspSerIleAspArgSerGluGluMet 1497
 DB 4873 TATTTCGAGAAATGAAACAGTGCACAGAAAGCTGACAGCGTACTCTGAGATGGAG 4932
 QY 1498 IleLysValLeuGlyGlnLysLeuGluArgAsnGlnTyrLeuGlu----- 1513
 DB 4933 TCCAGATTGGCGCGAGAAAGAAACAGACGACGACACTGCTGAGCTGGAATAGCA 4992
 QY 1514 ArgLeuGlnGluGluLysLeuGlnLeuSerAsnLysLeuGluIleLeuGlnLysGluMet 1533
 DB 4993 CGACTCCAGCTACAGAGCTGTGACTTAAGTCTCGG----- 5028
 QY 1534 GluThrSerValLeuLysAspAspLeuGlnGlnLysLeuGluSerLeuLeuSerGlu 1553
 DB 5029 ---TCTTGGCTGGATCGACACAGACAGATGCTATTCAGAGC----- 5067
 QY 1554 AsnIleIleLeuLysGluAsnIleAspThrThrLeuLysHisLysSerAspThrGlnAla 1573
 DB 5068 ---CGAAATGAGAGCTGTGACATCAAAAGAACATACTCTGAGAACTACAGAA 5118
 QY 1574 GlnLeuGlnLys-----ThrGlnGlnLeuGln 1583
 DB 5119 AGAACACCAAAAGCATGATGTCATCAGATTGTGATTAAGATGCTCAGCAGAGCTCAT 5178
 QY 1584 Leu-----AlaLysAsnLeuAlaIleAlaAspAspAspCysProIle 1598
 DB 5179 CTAGACATTCAGAAATATACAGACTGTGCTGATGAAACCAACAGAGAGTGTCTGG 5238
 QY 1599 ThrGlnLysGluThr-----SerIle 1606
 DB 5239 GAACAGTCCCGACATACCAATTATGAGCTCCAGGGGAAAGATAAACCCAGGGCTCTCA 5298
 QY 1607 AspCysValHisProLeu----- 1612
 DB 5299 GAATGCATTTTGAATTGTCTTTCTGCTCTAATGCTTTGGTACCTATGATTTCTGTG 5358
 QY 1613 ---GluGlnLysIleLeuLeuThrGlnLysGlnLysGlnLysGlnLysGlnLysGln 1630
 DB 5359 GGAATACAGAGAAATATCATCTTCACTCGGGGTAAAGAGACATCAATAGAAAT 5418
 QY 1631 GluLysLeuLeuHisGlnLysAsnGlnLeuGlnAlaGlnValGluLeuLysGln 1650
 DB 5419 TTGGATTACTTCATGTATGAGAGACCGTAGACAG-----AAA 5457
 QY 1651 ValIleHisLeuMetLysSerMetIleGlnSerLysSerSerLeuGlnLysLeuGlnHis 1670
 DB 5458 GTTAAAGTTTGTCTAATGAAATGAAAGATTAAGACTCAAACTC----- 5502
 QY 1671 GluLysHisAspThrGlnGlnLysLeuAla-----LeuLys 1683
 DB 5503 ---CATTTACAGAGAGTACACTAATGACCAAAATTTGAAGCTGATGAAATGGAA 5556
 QY 1684 GlnGlnMetGlnValValThrGlnGlnLysGlnLeuGlnGlnThrHisGlnHisLeu 1703
 DB 5557 AAAATAGTTGGGAACTTAAGAAAGAAACCTGATTTTAAGTGAAATAATGGAAATATTTT 5616

QY 1704 ThrAlaGlnValAspHisLeuLysGlnAsnIleGluLeu-----GlyLeuAsnPheLys 1721
 DB 5617 TCTGTGTATCCACCAGAGTACTCCAGAGATGAAACTTTCGAAGGCTCATATTTGTAT 5676
 QY 1722 AsnGlnAlaGln---GlnLysThrThrLysGlnGlnCysLysLeuAsnGlnLysLys--- 1739
 DB 5677 TTAGAAATGACATGCAATTAATCATACAGTGAAGATATTGGAGATATATGTGCCAAAGTG 5736
 QY 1740 -----GluLeuGlnLysSerGlnHisArgLeuGlnCys 1750
 DB 5737 AATGACACTGGAGAGAGATTTTGTATGTGAAATAATGAGCTGATGATGATGATGATG 5796
 QY 1751 Glu-----IleGlnLysLeuMetLysSerLeuLys 1760
 DB 5797 GAGAAAGCTAGCATTTGACATGAGACCCCTTACCTGAGGCTGACTTGAAGGTACTTCA 5856
 QY 1761 AspLysGlnSerValLeuLeuThrLeuLysGlnSerGlnLysValIle---AsnLeu 1779
 DB 5857 ACAGAGAAAGCTATTGTAAGAAAGAAATGAATGAAGAGATGATGATGATGATGATG 5916
 QY 1780 AsnGlnGlnMetGlnMetValMetLeuGlnMetGlnLysLeuLysAsnSerGlnArgThr 1799
 DB 5917 GAGAGAGAACTCTCAGTGTGTCACAGTGAGAGAAACAGCTTCTGAGAAATTAGATCT 5976
 QY 1800 ValIleAlaGluArg-----AspGlnLeuGlnAspAspLeuArgGlnSerValGlu 1816
 DB 5977 ATGTCAAAAAAACCCAGCGGCTGATCATGTTGTGTAATAATGAGAGAAACACAA 6036
 QY 1817 MetSerIleGluThr-----GlnAspAsp 1824
 DB 6037 ---GAGCTTGAATCTCATCAAAAGTGAATGTCATTCATTCAGGTGAGAGAGCAGAG 6093
 QY 1835 LeuArgLysAlaGlnGlnAlaLeuGlnGlnLysAspLysValGlnGluLeuThrSer 1844
 DB 6094 GTGAAGAGAAACAGCAAGCTCTCAGACTTGTCTGTGATGATGATGATGATGATG 6153
 QY 1845 GlnIleSerValLeuGlnGlnLysIleSerLeuLeuGlnLysGlnMetLysArgVal 1864
 DB 6154 GACAAATCTCATCTCCAGAGAAAGCTCCAGAGTTTGAAAGAGACTCAGAGCACTGTCT 6213
 QY 1865 AlaThrValLysGluThrLeuSerGluArgAspAspLeuAsnGlnSerGlnHisLeu 1884
 DB 6214 TTGACAAATGTGAGCTGGAACCAATTCACACATCAATTAAGAAAGAAATTCGTT 6273
 QY 1885 PheSerGlnIleGlnThrLeuSerLeuSerLeuLysGlnLysGluPhe----- 1900
 DB 6274 GTCAAGAAATGTGAAGCTCGAGGCCAGACTGATGATGATGATGATGATGATGATG 6333
 QY 1901 ---AlaLeuGlnGlnAlaGlnLysAspLysAlaAspAlaAlaArgLysThrIle 1917
 DB 6334 GTCTCCAAAGCTTGGAGGCCCACTGTGCGAGAAAGGTGATGATGATGATGATGATG 6393
 QY 1918 AspIleThrGlnLysIleSerAsnIle----- 1926
 DB 6394 TCACACAGAGAGAGATCATCAGCTGAGAGAGCATCGAGAACTGAGAGTTGCAAT 6453
 QY 1927 ---GluGlnGlnLeuLeuGlnGlnAlaThrAsnLeuLysGlnThrLeuLysGlu 1943
 DB 6454 GAGCCCGATGAAGAAAGACAGCTGACATCGAGAGAAATGAAAGAAAGCCGAGCGGAG 6513
 QY 1944 ArgGlnSerLeuIleGlnCysLysGlnLysLeuAlaLeuAsnThrGlnHisLeuArgLys 1963
 DB 6514 AATGATTCACAT---AAGATTAAGTGTGAAGACCTTGAAGGAATTCAGAGATG 6564
 QY 1964 ThrLeuLysSerLysAspLeuAlaLeuGlnLysMetGlnGlnLysArgAspGlnAla 1983
 DB 6565 TCAGAGAGAAACAGAGAGCTGATGAT---CTTGATGCCGAGAAATTCACAAACA--- 6615
 QY 1984 AsnLysValIleAlaLeuThrGlnLysMetSerSerLeuGlnGlnLysAsn 2003
 DB 6616 ---GAGGTAGAGACTTAAAGAACCAAAATGAGAAAGATG 6651
 QY 2004 ValThrThrLeuLysGlnGlnLys-----GlyLysLysGlnThr 2016

```

Db 6652 GCCAGAAAGCCGTAAGTTTGGATTAGACCTTGTACGTTAAGCTCGTGAAGAAAT 6711
Oy 2017 PhetylLeuGlnArgProSerLysGlnGlnSerSerSerGlnMetGlnGluLeuArgGlu 2036
Db 6712 CTGACAAAACAAATACAGAAAAACAGGTCAGTTGTCAGACTAGCAAGTTACTCTCT 6771
Oy 2037 SerLeuLysThrLysAspLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 2056
Db 6772 TCATTAAAGTCTGTAGAGAAAGAGACCAAGCAGATACAGTACCAAGAAATCT 6831
Oy 2057 -----AsnGlnLysAsnLeuThrAlaLysLysSerSer 2068
Db 6832 AAAATGCACTGAGAGATGCTTCAAGATCAGTTAAAGAGCTAATGAGGAGAGAGGCC 6891
Oy 2069 Leu-----GluGlnGlnLysLeuGln-----AsnAlaSerLysLeuGlnAlaValSer 2085
Db 6892 TTGGTGTGACCAAGAAATATATGAAAGCCAGACAGAGCTAGACCAACCAATGAG 6951
Oy 2086 GluArgGlnAsnLeuArgHisSerLysGlnGlnLeuValSerGlnLeuGln----- 2102
Db 6952 GAAGAGCATGCTGAGAAATAGATGAAAGCTGAGAGCCGCCCTAGAAAGCTGATGAA 7011
Oy 2103 -----GlnLeuSerLeuThrLeuLysSerArgAspHisAlaPheAlaGlnSerLysArg 2120
Db 7012 AAGAAAGCACTGTGTGCTTACAAACACTGAG-----GAAAGTGACAT 7056
Oy 2121 GlnLysAspGlnAlaValAsnLysLysLysSerLeuAlaGlnGlnGlnGlnGlnGlnGln 2140
Db 7057 CATGCAGATTACTTAAAGGTATAGGTAGAGACACTTCAAAAGAGCTAGAGATGCCAG 7116
Oy 2141 LysGlnMetAsp-----GluPheArgAspSerLys-----GluSerLeu 2153
Db 7117 ACAAAACCAAGACATGACGCTTGGAGCAGAGAAATCCAAAGAGAGATGAGACCCTA 7176
Oy 2154 GlnGlnGlnSerSerHisLysSerSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 2173
Db 7177 AAGCAAAAATAGAGGATGACCCCAAGTCTGAGAGCTGGAATTAAGTGTGTTACT 7236
Oy 2174 LeuLysGlnGlnLysGlnLysAspLysAsnLysLeuAlaGlnLysValLysGlnValAsp 2193
Db 7237 ATAAAGTCACAAAAGAAATCTGACAAATGCAATTAACAAAAGAGCAAGCCAAATCT 7296
Oy 2194 GlnLeuLeuGlnHisLysSerSerLeuLysGlnGlnGlnGlnGlnGlnGlnGlnGln 2213
Db 7297 GAATTA-----GAAATA 7308
Oy 2214 ArgAsnGlnLysLeuArgAsnTyrGlnLeuGlnGlnLysMetAspLysMetGlnLysGln 2233
Db 7309 ATAAATTCATCATTTGAAAT----- 7329
Oy 2234 LysSerValLeuArgLeuMetGlnAsnGlnProGlnGlnGlnGlnGlnGlnGlnGlnGln 2253
Db 7330 -----ATTTGCAGAAAAAGAGCAAGAGAAATGACATGAAGAAAAATCAGACAT 7383
Oy 2254 ArgMetAspLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 2273
Db 7384 GCCATGAGATGCTTCAAAAG-----CAATTAAGAGCTCATGACAGAGTGGCA 7434
Oy 2274 AlaValTyrSerGlnGlnHisThrLeuLeuSerSerLeuSerSerGlnGlnGlnGlnGln 2293
Db 7435 GCCCTGCAATATGACCA----- 7452
Oy 2294 ThrGlnAlaHisLysHisCysMetLeuAsnLysGlnSerLeuSerSerThrLeuSer 2313
Db 7453 -----GAAGCCTGTAG-----GCCAAAAGCAAGAACTTAACTAG 7485
Oy 2314 ArgSerPheGlnSerLeuGlnThrGlnHisValLysLeuAsnThrGlnGlnGlnThrLeu 2333
Db 7486 AGTCAAGTAGAGTGTCTTGAAGAGGCTGAGTGTGACAGAGCCCTGATGAGGCC 7545
Oy 2334 LeuAsnLysPheLysValValTyrArgThrAlaAlaValLysGlnLysPheHisSerLeu 2353
Db 2334 LysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 2353

```

```

Db 7546 AAAAATATATATGTTTTCGAATCTTCACTG-----AATGCCCTCAT 7590
Oy 2354 LysAspTyrGlnLysAspLeuAlaAlaGlnGlnLysArgHisAspGln-----Leu 2370
Db 7591 CAAGAACTAGAACTGGCAAGCAGAACTGAGAGAAAGAGATGAGAAATCAGTACTG 7650
Oy 2371 ArgLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 2390
Db 7651 AAAAATCAAAATCAAGACCAAGAGCAGCTGTGCTTAAGTCTCCAGGTGGAAGAGAG 7710
Oy 2391 GlnLeuLysPhe-----CysGlnLysLeuPheLeuAsnGlnLeuLeuPheLys 2406
Db 7711 CACCAACTTTGGAGAGAGCAAACTTAGAACTGAGAAATGTACAGATGAAATGGACAG 7770
Oy 2407 LysAlaAsnLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 2426
Db 7771 AAGATCCAACTGCTACAAATCCAAATGCTCTTGGAGAGACATTAAGAGTGTCCAG 7830
Oy 2427. GlnValGlnSerThrLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 2445
Db 7831 AGTTCTTACAAAGATCTAGAGAAATGAGCTGAAATTCACAAAATGACAAAATGTCT 7890
Oy 2446 LeuGlnGlnLysPheLys-----AspLeuHisVal 2454
Db 7891 GTTGAAGAAATCAAAATAGCTGCAAGAACTGAGCTGAGAGAGGAAATGACATGAG 7950
Oy 2455 AspAlaLysLys-----LeuSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 2471
Db 7951 ATGGCAGCAAGAAACAGCAGAGCTGCAAGAAATGCAAGTACAGAGAAATATAGCT 8010
Oy 2472 SerThrLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 2491
Db 8011 GGAGAGTGTGAGTACTGTTGGAAGAAATGAAGAGC-----AGCAAGATCA 8058
Oy 2492 GlnLeuThrValTyrLeuAsnGlnPheGlnAlaLysLeuGlnGlnGlnGlnGlnGln 2511
Db 8059 -----TTGAAGAGCTCACATGAAAT 8082
Oy 2512 LysGlnLeuMetArgArgMetGlnHisLysGlnProSerLysSerValMetGlnGlnGln 2531
Db 8083 AGTGAATTTGAAGAG-----AGCTACATTTGCAAGCAAGCAAGCAAG 8121
Oy 2532 AsnAlaArgLeuGlnGlnLysLysLysLysLysLysLysLysLysLysLysLysLys 2551
Db 8122 CAGCTGCAAGAAAGAGAG-----AAAGTGAAGAGAAATAGCTGATATCAGCTA 8172
Oy 2552 ArgLysMetLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 2571
Db 8173 CGG-----CTTCATGAAAGCT 8187
Oy 2572 GlnLys-----ValAlaLysLeuGlnAspLysLeuSerArgAsnAlaGlnAlaGln 2589
Db 8188 GAAAGAAACACAGAGCTTGTGTTGAG-----ACAAACAAAGATGAAGTGAAGAA 8241
Oy 2590 LeuAsnAlaMetGlnValLysLeuThrLysLysGlnAspAsnLeuGlnAlaAlaMetLys 2609
Db 8242 ATCCAGACATACCGAAGAAATGACTTTAAAGAAAGATGCTCAGTTACAGAAAGCTG 8301
Oy 2610 GlnLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 2629
Db 8302 GAGATGAGCTTTAAAGTCTAGT-----AAAGAGAGCTC 8337
Oy 2630 AspAsn-----LeuLysThrLysValValLysLysLysLysLysLysLysLysLys 2642
Db 8338 AATTAATTCATTTGAAGCTACTACTACTAGATTTTGAAGATTTGAAGAAACCAAGATG 8397
Oy 2643 LysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 2654
Db 8398 AATCTAAATATATTAATAGTTGAAGAGAAATGACAGCTGCCAGGAGAAATGAGAG 8457
Oy 2655 AlaTyrLeuLysSerLys-----LeuGlnAspLysGlnGlnGlnGlnGlnGlnGln 2672
Db 8458 TTGTTGATCAAAATCTGTAAACAGCTGGAAGAGGAAAGAGATTAAGTCAAGAAAGATC 8517

```


Db 372 AGCACCACGATTTGAAATTTGATGTAATCAACACAGTGTAAATGAAATGACATTAATAGTTC 431
 Oy 220 TGLAASCyshepGLValAlaMetValSerHisLeuAsnLeuValAspLeuAlaGlyse 240
 Db 432 TCACAGGAGTGAATCAACCACTGTCTGTTCTCGTAATCTCAATTAATGAGAACTGACATAGTGG 491
 Oy 240 TGLValGAlaSerGlnThrGlyAlaGluGlyAlaArgLeuLys 254
 Db 492 AGAATTAACAGATCATGATGACGAGGGCTTCTGTGGACGCGAAAGTGAATTCACAC 551
 Oy 255 ---GluGlyAsn---LysAspSerLeuPheLeuGlyGlnValIleGly 271
 Db 552 ACCGATGACTGCTCACTTCACAGAGTAATGTGGTCACTTTGTGATG 597
 Oy 271 slyLeuSerAspGlyGlnAlaGlyGlyPheIleAsnTyArgAspSerLeuThr 291
 Db 598 ---AGAAGCAGAAACCTACAAA 617
 Oy 291 gtleLeuGlnAnserLeuGlyGlyAsnAlaLysThrValIleLeuGlyThrIleThrPr 311
 Db 618 TTTATTAAAGGAAGAAGATTGGT 642
 Oy 311 ovalSerPheAspGlyThrLeuSerThrLeuGlnPheAlaSerThrAlaLysHisValArg 331
 Db 643 ---GTTGAAGATTCTTATCT 674
 Oy 331 gasThrProHisValAsnGlyValLeuAspAspGlyAlaLeuLeuLysArgTyArgIle 351
 Db 675 AGACAGTCCGCACTATCTAGAGATGATGGAAAGTGAAGTGGCTGGAGACAGCAT 729
 Oy 351 sglIleLeuLeuSerLysLysGlnLeuIleAsnLeuGlnSerSerGlyThrIleVal 371
 Db 730 -GAGATTGAAGAGTAAACAGAGAGCTGGAGAAAGAGAGAGCTTACCTATGGGACATGGAAG 788
 Oy 371 a-----GlnAlaMetAlaLysGlnGlnHis-----ThrG 381
 Db 789 ACTGCAGACAGTTACAAAGATTGAAAGCTGCCATTAAACAAAGAGATGSCATTATTAACCA 848
 Oy 381 nleuLeuAlaGlyIleLysGlnLeuHisLysGlnValArgGlnAspArgIleThrPheLysLeu 401
 Db 849 GCTCAGCTGTATTATTCACAAACAGAAAGAGAAAGAGATGAGACAAATGACAGAAATTGTT 908
 Oy 401 rsnIleValAlaLaserSerGlnLysSerGlnLysPheGlnValValLysArgLys 421
 Db 909 AGAGTGG-----ACAGACAGAGTCAAAAAATTTACAGATTCAT 947
 Oy 421 gatValThrThrAlaArgGlyLysIleGlnAsnSerLeuHisAlaSerGlyValSer 441
 Db 948 TCAGCAATTCACAGAGTATGTAACCTGTGAGAAACAGACATCATATGTGTGACACAGCTCA 1007
 Oy 441 P-----PheAspMetLeuSerArg 447
 Db 1008 CTTTACTCAAGCCAAACACAGATCTCTCATCTATCACAAGCAGCTTGAGAACAGACGCA 1067
 Oy 447 gLeuProGlyAsnPheSerLys 454
 Db 1068 CTTTATTAGCAATTTATCAGAAAAAGAAAGAAAGACCTTCACAAATGCAATTTAGTTCTTGC 1127
 Oy 455 ---LysAlaLysPheSerAspMetProSerThrPheProGlyIleAspAspSerValCys 473
 Db 1128 AGAGAAATTTTAAAGATATGCAAAAGAACACAGATTAATAAAAGTGAACAACTCAATTAAGA 1187
 Oy 473 rGluPheSerAspPheAspArgAlaLeuSerMetMetLysPheSerAsnGlyIleAspAlaG 493
 Db 1188 AGAATATACAGAAAGAGACAAACATCTTGAAGATTAACACAAATAATATATAGAGAAAGA 1247
 Oy 493 u-----TrpAsnLeuAlaSerLysValThrHisArgGlyLysThr-----SerLe 508
 Db 1248 AAGAAACCTGTGAGCTTAAGATATTTAAACAACTGCGATTAATTTACTAGAGAAAT 1307
 Oy 508 uHisGlnSerMetIleAspPheGlyGln-----IleSerAs 520
 Db 1308 ACAAGAAACAGATTGTGCACAAAGACCAAGAAATTAATAAAACATGAAATTTAGAGCTGACTTA 1367

520 pSerValGlnPheHisaspSerSerLySLuasnGlnLeuGlnIleuProLysaspse 540
 1368 TTTCAAGCAAAAACAAAGACACGCTTCTCAAGAAATTAACAGCTTAAGG----- 1416
 540 rGlyAspMetAlaGluCysArgLys-----AlaSerPheGluLySGluI 555
 1417 -GGGCACTGCCAAGACACTCGAAGAGCAAAATCATTAAGACACGCCACTGTGTAAT 1475
 555 ethrSer---LeuGlnGlnGlnLeuGlnSerLys---GluGluGluLySLySGluLeu 573
 1476 ACTACACAGCATGAGACAAACAAACACAAAGAAAGTTAGAAACAATCTCCGGCGAGCTGCA 1535
 573 LginSerPheGluLeuLySLeAlaGluLeuGluGlnLeu-----SerValLySAI 591
 1536 TGGAGTGTATGGGCAACGATATAGTCCAAATGCAATGAACAGATTTATAGAACACACATGGC 1595
 591 aLysAsnLeuGluMetValThrAsnSerArg-----GlnHisSerLysAlaAla 608
 1596 ACAGATGAGAGAAATGAATGAACACGGCATTAACGGAGAAATGACAGAAATGCTTTAGGCATTA 1655
 608 uValGlnThrAspValGluLySLyGluValValArgLysGluMetSerValLeuGlnLysPse 628
 1656 TTCAATATTTACGATTAAAGACATGACATTAAG-----TTATATGAAATGT 1700
 628 rGlyTyAsnAlaSerAsnSerAspLeuGlnLysPseSerValaspLySLyArgLeu 648
 1701 GGCATATAATGACATGTAATTAATAATTTGACAAATCTCAACTCAACCTCAAAAGCAAACTCAA 1760
 648 rSerSerHisaspLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 668
 1761 GGAGAACTACAGACTAATTTATTAAGAAAGAAAGTGTGCTCTACAGACACAGCTTGAAGACCT 1820
 668 uGluGluPheIleGluLysLeuLysLysSerLysSerGluLysAsnLysLysLysLysLys 688
 1821 TGTTAAGAAATTTAGAGCTTTTCAAGAGAAACAGATTCAAGACGTACAGACATAGCTGCA 1880
 688 uGlnAspPheMet-----GlnSerIleGlnLeuCysGluAlaIleMetAl 703
 1881 ACAGAAAGTAATTAATTAATGAAGACATTAATGCTCTAGTACAGTGGAGATTTGAAAGC 1940
 703 aGlu----- 704
 1941 TGAATTTGTTTCTGCATCTGATTCGATTCGAAAGAAAGAACTAGAAATTAACATGAAAGCAGAA 2000
 705 -----LysAlaAsnAlaLeu----- 709
 2001 TACAAATTACAGATTAACCTTGAAATGTTAGAAAAGAAAGAAAGTCTGTGTAGACAG 2060
 709 ----- 709
 2061 AATGGCTGATTCACAAGAGCTGAAATTAGAGAGCTGGAAACACAGCTTATTATTAGTCA 2120
 710 -----GluGluLeuAlaLeuMetArgAspAsnPheAsp-----AsnIleI 723
 2121 CGAAAGAAAGAGCTTTCCAAACTGAAAGAAAGATTTAGAAATTTGAAACATCGATTAATATTGA 2180
 723 eLeuGluLysGlnIleuLeu-----LysArgGlnIleAlaAspLeuArgse 739
 2181 AAAACTTAAAGTAAATTAATTTAGCATCTATTAACACAGCATAGATGCTTTACAGATGA 2240
 739 rLeuLySGluLysGlnGlnIleuThrAsnGluPheGluIleLeuGluLySLyGluThrGlnLySGl 759
 2241 AATGAGTCAAAAGATGAAAGAACATCGACAGTT-----GAAAGAGCAATTTGATATAC 2291
 759 uHisGluAlaGlnLeuIleHisGluIleGlySerLeuLySLyLeuValaGluAsnAlaGlu 779
 2292 TAAGCAGAAATCAATTAATTTGGAAATTTCCAAAGCTTAAGCATTTACACAGCTCT----- 2346
 779 uMetTyAsnGlnAsnLeuGluLy-----AspLeuGlnIleThrLysTh 793
 2347 -CTTGAAATTCAAAGTCAAGAAAGATGACTCTTCAATCAATCAAGTCACTTCAAAAAGAAAT 2405

QY 793 rlyseuleuysglu-----GlnGlu1 801
 Db 2406 TGAATACTCAGACAGAAAGAAAGAGGTACTGTACAGAAAGTCAAGAAATT 2465
 QY 801 egluleu-----AlaGluLeuArgLysArgAlaAspAsnLeuGlu1 815
 Db 2466 ACAACTTAAACAGAAATTGTAGAAAAACAGATGAAAGAAAAAGAAATGATCTTCAAGA 2525
 QY 815 slyValArgAsnPheAspLeuSerValSer---MetGlyAspSerGluLysLeuGly 834
 Db 2526 AAAATTTCACCAACTTGAAGCAGAAATGACCTTCTTAAAGTGAAGAAAAAGCCCTTGA 2585
 QY 834 ugiuilepheGlnLeuLysGlnSerLeuSerAspAlaGluAlaVal----- 849
 Db 2586 AGACATGTGAAGAAATACATCTCCGTAGCCACAGAAAGAAATGATTTCTTAGACTC 2645
 QY 850 -----ThrArgAspAla-----GlnLysGluCysSerPheLeuArgSerG1 863
 Db 2646 CATTAAGTCCAAATCCAAAGACTGTGTGGAAAAAGAAATGAAATACCTTATAGAGA 2705
 QY 863 uasn-----LeuGluLeuLysGluLysMetGluAspThrSerAs 876
 Db 2706 AAATGAGACCTCAACAAACAACTGATTACGCTAATATGAAAGATTGAAAGCAAGGA 2765
 QY 876 ntrPtyrAsnGlnLysGluLysAlaAlaSerLeuPheGlnLysGlnLeuThrGlu1 896
 Db 2766 CACTTTTCATTTGCTGAAAAAACTTGAAGTAACTATCAAGAGTTACAAAGAGAGA 2825
 QY 896 sSerAsnTrpLysLysMetGluAlaAspLeuGln-----LysGluLeuG1 911
 Db 2826 TGCCTGCTTCAAAGTAAAGATTTAGAACACATAAATAAAGAGATTTAGA 2885
 QY 911 n-----SerAlaPheAsnGluLeuAsnTrpLeuAsn----- 921
 Db 2886 GTATTAAGTAACTTAAAGCCTTAAATGAAGAGCTTCATTGCAAGAAATTAATCCAAC 2945
 QY 921 ----- 921
 Db 2946 TACAGTAAATGAAAGTCTGCTTGTGATGAAGACAAACTTTGTAGCAGAAACATT 3005
 QY 922 -----GlyLeuLeuAlaGlyLysValProArgAspLeu----- 933
 Db 3006 GGAATGGGTGAGGTGTGGAAGAGATCAACAACTCATGGAATAACTTGAGGTAAAC 3065
 QY 934 -----SerArgValGluLeuGlnLysLysValSerGluPheSerGlnLeuGlu1 951
 Db 3066 CAAGCAGAGAGAAATTAGAGCTGTACAGAGACTGTCTGATCTTTCTGCAACATTTGAAACA 3125
 QY 951 salaleuGlnLysAsnAlaLeuGluAsnGluValThrCysLeuSerGluTrpLys-- 970
 Db 3126 GAACATGGTGAGATTAGTTTCTTAATGAAGAAAGTTAAATCTTTAAAGCAAGAAAGA 3185
 QY 971 -----PheLeuProAsnGluValGluCysLeuLysAsn----- 981
 Db 3186 ACAAGTTTCATTGAGATGTAGAGAGCTGAAGATCATTTATTAACCCACAAGGCGCAAGAA 3245
 QY 982 -----GlnLysSerLysAlaSerGlnGluLeuMetLeuLeuLysG1 995
 Db 3246 TGTACAGTCATGTACTACTAGTAAGTCTTTTATTAGATGAGATTGTGACCATGACAAG 3305
 QY 995 nglucly--GluHisSerAlaSerLeuLys----- 1006
 Db 3306 CAGGGTGCTGAGAGATCAGTTCTTAAGTAATTAAGTTTGGTGAAGATCAAAAT 3365
 QY 1007 -----GlnGluLeuLeuMetGlnGlnSerGluGln-- 1017
 Db 3366 AATGCTGAGATAAAGTTCTTTTGAAATAATGATGCTGGAGAGAAATAGACACAGA 3425
 QY 1018 -----IleLeuGlnLeuThrAspGluValThrHisThrGlnSer-----LysValG1 1033
 Db 3426 ACAGTTGATTTTGGATCTTACCATCTGTACAAAGAAATCATTACTTAGAGCAACTCA 3485
 QY 1033 nGlnThrGlnGluGlnTrpLeuGlu-----MetLysLysMetHisAspAs 1048

Db 3486 ACCAGTGAATATGATTAATCACTCAGAAAGAACTCAATGTACTTAATTCAGAAAGAAATGA 3545
 QY 1048 pleuPheGluLysTrpIleArgAsnLysSerGluAlaGlu----- 1061
 Db 3546 TTTA-----AGCTTACAGATGGAACCCACCATTTGCTCTCTGCT 3590
 QY 1062 -----AspLeuLeuArgGlu---MetGluAsnLeuLysGlyThr----- 1073
 Db 3591 TTATTCAACTATGTGTGATCAGGTTCTGTAAATATGAAAAATGAAGATTAAGCTCT 3650
 QY 1074 -----MetGluSerValGluValLysIleAlaAspTh 1084
 Db 3651 TTGCAGTCTTAAGAAAGACTATTTTGTCTCAAGAGAAAGATCAAGAACTTCACAGA 3710
 QY 1084 rlyshLsgluLeuGlu---GluThrIleArg-----AspLysGluGlnLe 1098
 Db 3711 AATACACAGTTAGAACTACAGACTATGAAACACAGAAACAGCTGATGAAGAAACC 3770
 QY 1098 uLeuHis-----GluLysLysTrpPhePh 1106
 Db 3771 TTTCATCTGCTCATTTGAAAACTTCAAAAGCACTGTCTGAAGATGTTCTTATTTT 3830
 QY 1106 eGlnAlaMetGlnThrIlePheProIleThrProLeuSerAspSerLeuProPheSerLys 1126
 Db 3831 ACAGACTTTATGCACTGTC-----CTTGATGAATATTAATCTCTGCTT 3875
 QY 1126 sleuValGluGlnAsnSerGlnAspProIleGluLeuAsnAspTrp----- 1141
 Db 3876 AAATGTGAAGTAAATGCAAGAAAGCAAGAAAGAAATTCGTGTATTCATTTCTGAAATGA 3935
 QY 1142 -----HisAs 1143
 Db 3936 AGATCCAGAAATTCAAGATTATAGATATGAACTCAAGACTTCAAGAAATATATCACAC 3995
 QY 1143 nLeuIleAlaLeuAlaThrGluArgAsnAsnIleMetValCysLeuGluThr----- 1160
 Db 3996 TCTTCTCAACAAAGTAAACAGAAATATACAAACAACTTGTGACTTCAACACGACTAAG 4055
 QY 1161 -----GluArgAsnSerLe 1165
 Db 4056 CAAGATCTGGGCGACAGACAGATGATGAACTGAATTTGGAGAAACAACTTCC 4115
 QY 1165 uLysGluGlnVal-----IleAs 1171
 Db 4116 AAAGAGGAAACAGAGTTTATATCAATCTTCAATGACCAATTGGAAGACATTGA 4175
 QY 1171 pleuAsnThrGlnLeuGlnSerLeuGlnAlaGlnSerIleGluLysSerAspLeuGlnLys 1191
 Db 4176 TGTCAATCATTAAGCAAGATTATCTTCTGCAACATCTTGAAGAAACTTAAGTGAAGA 4235
 QY 1191 sProLysGlnAspLeuGlu-----GluGlyG1 1200
 Db 4236 ACAAGTTCAAGAAATTAAGAGCCATCATCTTCTTTCAGCAACAGTGAAGAAACATGA 4295
 QY 1200 uValLysLeuLeuLeuGluMetGluLeuLeuLysGlyHisLeu----- 1214
 Db 4296 ACAAACTATGAGCGCAGAGATCCACTGTTTACAGAGAGGCTTCAAGCTGTTAGTACTC 4355
 QY 1215 -----ThrAspSerGlnLe 1219
 Db 4356 CACGGTCCGCCAAGCTTACCTGTTGATTCGGTGTATTTACAGAAATGTATGCCACNAG 4415
 QY 1219 uSerIle----- 1221
 Db 4416 AACCATGTACCCCTGGAAGTTGTGAAAAAGAAATATGATGTATCAATATAGATTTCTGG 4475
 QY 1222 -----GluLysLeuGlnLeuGlnAsnLeuGluValThrGluLysLeuG1 1236
 Db 4476 TGAATTTGAGTGAAGAGGAAACAAATATATGCTTAACTTGTGAAAAACAAATNCCAAAGA 4535
 QY 1236 nThrLeuGlnGluGluMetLysAsnIleThrIle-----GluArg 1249

```

Db 4536 ACAATTAGAAAGAGTAGCTAAGTTATGTGTCAATGATATAGCATTTGCTCACA 4595
QY 1249 GAANGUleu----- 1252
Db 4596 AACGGAACGTCTGATATCTGGGGAAAAGAAATAGTCATCATCAACAGCACA 4655
QY 1253 -----GlnThrAsnPhgLeuAspLeuLysAlaGlnHisAspSerLe 1266
Db 4656 TGCTGTGTGTACAGCAAGAACATTTATTTATGAAATGAAATTTTCACAGATCAAT 4715
QY 1266 uLysGlnAspLeuSerGlnuSnllleglInserlleglUthrGlnAspGluLeuArgAl 1286
Db 4716 TGGTTTCAGACTTTTGACAGACGGAT-----GTCAATTTAAAGAGAAATTTAAAC 4769
QY 1286 aalaglIngluLeuArggluIn---LysGlnLeuValAspSerPheArgGln----- 1303
Db 4770 ACTTAGTAAAGAGTTAGAGAACATGGAAGAAATTTTATTCAAATGATGATCCCA 4829
QY 1304 -----GlnLeuLeuAspCysSerValGlyIleSer-----SerProAs 1316
Db 4830 TGATATACCAATCAAAAGACGTGTGTGCTGACTATTTCAAGAAATGTTCTCCAAAG 4889
QY 1316 nHisAspAlaValAlaAsnGln-----GlnLysValSerleuGluValAlaAsnSe 1333
Db 4890 TAAACATTTATAGTTAGACAGTCTATTCATGATGAGATTCAGTGTCAAGCATGATGC 4949
QY 1333 r-----LeuGlnSerGluMetLeuArgGlyLysIuArgAspGluLeuGlnTh 1348
Db 4950 TTCTAGACAACTAATCTTGAATGAAAGACAGTTGCAAGATGAGACAGCACTTGACG 5009
QY 1348 rSerCysLysAlaLeuValSerGluLeuGluLeuArg---AlaHisValLysSerVa 1367
Db 5010 ACAATACCAAGAACATCAACAGCAGCAAGCAATTTGTTAAGCAGCACATATGCCGCAAT 5069
QY 1367 lGluGluLysuSnllleglUthrLysLeuAsnGlnLysLeuGluLysGluIleLeuG 1387
Db 5070 GCGAG----- 5073
QY 1387 yLysSerGluGluSerGluValLeuLysSerMetLeuGlnuSnllleglUthrGluAspAsnAs 1407
Db 5074 -AGACAGCGAGAAAGACCAG-----GAAACAGCTTACAAAGAAAGCATTA 5114
QY 1407 nLysLeuLysGluGlnAlaGluLysuSerSerLysGluAsnGlnPheSerleuGluG 1427
Db 5115 GAGCTTAATAGACAAATTAGCCAGAGATCCCTCATAGTAATGAACCTGTTTCAGA 5174
QY 1427 uValPheSerGlySerGlnLysLeuValAspGluIleGluValLeuLysAlaGlnLeu 1447
Db 5175 G-----AGAGAGAGGGTCTTTAGAGAGCTGGAGACCTAAACACAGCTGCTTT 5225
QY 1447 sAlaAlaGluGluLysLeu-----GluIleLysAspArgAspTyrPheGluLeuValG 1465
Db 5226 AGCTGGAAGAGAGAGAGCTGTGTGTGTGAGCTGCCACACAGCAGT-----ACCGA 5273
QY 1465 nThrAlaAsnThrAsnLeuValGluGlyLysLeuGluThrProLeuGlnAlaAspHisG 1485
Db 5274 AACACAGAAATGGAATGAAAACCAAGAGAGAGTTGAA-----GAACAAACATTTA 5324
QY 1485 uGluAspSerLeuAspArgArgSerGluGluMetGluIleLysValLeuGlyLysLe 1505
Db 5325 AGAAAAAGAAATTTAGCAAAAAACCTGAAAGATGTGCTCCGAGATTTTGTCTAATGA 5382
QY 1505 uGluArgAsnGlnUthrLeuGluGluArgLeuGlnGluLysLeuGluLeuSerAsn 1525
Db 5383 -----AGCTATGCACTCCAGAAAGCTAAT-----ATATAGACT 5414
QY 1525 sLeuGluIleLeuGlnLysGluMetGluThrSerValLeuLeuLysAsp 1542
Db 5415 TTTGAAGATCCCTTAGAGAGTTGAAGACAAACAGCACTGTTAAGAAACAATGTGTCG 5474
QY 1543 -----LeuGlnGlnLysLeuGlnSerLeuLeuSerGluAsnIleIleLe 1557
Db 5475 CCATGTCCTTGGGATTTCTAGATAGATCTAGTAAAGCAAGCATCTGCGCACTAATTTG 5534
QY 1557 uLysGluAsnIleAspThrThrLeuLys-----HisHisSerAspThrGlnAlaG 1574
Db 5535 GAGGTCAAGAGCAGAGCAGCATGTAAAGTCATGTCTCATGAGGAACATACAAAGATAC 5594
QY 1574 nLeuGlnLys----- 1577
Db 5595 AGATGAATCCATTCCTCTTATCTGTGAAGTGAATGCCAAGAAATGACATTAAAGATG 5654
QY 1578 -----ThrGlnGlnLeuGlnIleuAlaLysAsnLeu-----AlaIleAl 1591
Db 5655 GTCAAAAGTAAGTGAAGAGAGAGAGAGAGAGCTGTCAACAGCATTTGTAGAGAGGTTTGC 5714
QY 1591 aalSerAspAsnGlyProIleThrGlnGlu----- 1601
Db 5715 TGAACCTGAATTAACCCCTGAATTAAGAAACTATATGCTGAACATTAGCTCTGACTACA 5774
QY 1602 -----LysGlnThrSerAlaAspCysValHisPr 1611
Db 5775 AGCAGCACTGAAAAAATCCTTAGAAGCCATTAAGTGAAGTGAAGTGAAGTGAAGTGA 5834
QY 1611 olauGluGluLysIleLeuLeuLeuThrGluGluLysGlnLysThrAsnGluGln 1631
Db 5835 GAAAGTGAACAGACAGAGATTGATGCTGATCATTTAGACAGAAACAGAACAGCAGAGA 5894
QY 1631 uLysLeu-----LeuHisGluLysAsnGluLeu---G 1641
Db 5895 GTCCCTTAGTGCCTCAAGAGAACTTCGAGAGCCGCTTATGAGAGTCCAGGCGCAGAGA 5954
QY 1641 uGlnAlaGlnValGluLeuLysCysGluValGlnHisLeuMetLysSerMetIleGlu 1661
Db 5955 ACAGTAGCTGTGAGCTCAGT---AAGGTCAGAGGCGCTCATGTATGCGTATGCAATGA 6011
QY 1661 rLysSerSerLeuGluSerLeuGlnHisGlnLysHisAsp-----ThrGluG 1677
Db 6012 AAAAAGCTTTTGAAGAGCAAAATTCAGCAAAAATACTATATATAGTACCTTTGAGCA 6071
QY 1677 nGlnLeuLeuAlaLeuLysGlnIleMetGlnValValThrGlnGluLysGluLeuG 1697
Db 6072 GCGATTGTATGTGCAAGTAAACAGGTTCAGAAATTTGAGCGCAGCAAGCAGATCCA 6131
QY 1697 nGlnThrHisGlnHisLeuThrAlaGluValAspHisLeuLysGluAsnIleGluG 1717
Db 6132 AGAAGAAAGAAATTTACTGTCCAGACAAAAGAAAGAAAGTATGAAA-----GCAAGAGCAG 6185
QY 1717 yLeuAsnPhyLysAsnGlnAlaGlnGlnLysThrThrLysGlnGlnCysLeuLeuAsn 1737
Db 6186 C-----CCAGTTGAACAAACATTTACTACAGGA 6212
QY 1737 uAsnLysGluLeuGlnSerGlnHisArgLeuGlnGlnGluIleGlnGluMetLys 1757
Db 6213 GACAGAAAAATTAATGAAGAAACCAAGAGTAAATGCAAGCTGAAACAAAGAAAGTAC 6272
QY 1757 sSerLeuLysAspLysGluSerAlaLeuGlnUthrLeuLysGluSerGluGlnLysVal 1777
Db 6273 TGACCTTCAAAAACAAAGTGAAGCTTAGAA-----AT 6305
QY 1777 easnLeuAsnGlnLysGluMetGluMetValMetLeuGluMetGluLysAsnSerG 1797
Db 6306 AGATGTGAAGAAACAAAGTCACTAGG---TTTATGAGCTGGAACAGAAAAAATCTGA 6362
QY 1797 nArgThrValIleAlaGluArgAspGlnLeuGlnAspLeuArgGluSerValGlu 1817
Db 6363 ACTAATGATTTAAGACGCAAAACCAAGCACTTGGAAGAACAGTTCAGAAAAAATGAGAA 6422
QY 1817 tSerIleGluUthrGlnAspAspLeuArgLysAlaGlnGlnAlaLeuGlnGlnGlnLys 1837
Db 6423 ATTTTATGATGCAAGCATTTGACAGACATGAGAGAGATGTATTCACAG----- 6477
QY 1837 pLysValGlnLeuUthrSerGlnIleSerValLeu----- 1849
Db 6478 -GAAATACAGAAACATTAAGCAAGCACTTAAGTGTGTTCTCGATTCCAGCTACAGTGA 6536

```

QY 1850 -----GlnGluLysIleSerLeuLeuGluAsnGlnMetLeuTyrAsnValAlaThrVa 1867
Db 6537 ACATCAAACTAGAGAGGTTGAACAGTTAGCAAAATCATCTGAAGAAAAACAGCAAAATG 6596
QY 1867 llyGluThrLeuSerGluIleAspPheLeuAsnGlnSerIleGlnIlePheSerGln 1887
Db 6597 CAGTACCTTTTGTCTCTAAAGAGCAGCTTCAAGGGATATACAGAAAGAAAGTAAGA 6656
QY 1887 uileGluThrLeuSerLeuSerLeuLysGluLysGlnPheLeu----- 1902
Db 6657 AATGAGAACTGAGATTACAGATGAGAACTGAGAGCGCGCTTCTTGAGAGGACCG 6716
QY 1903 -----GluGlnAlaGluLysAspLysAlaAspAl 1912
Db 6717 AAACACTTTGGAGCTGAGAAAGCTAAACAGAAATTTGCTCCTAGACAGTAAATGACAGC 6776
QY 1912 alaArgLysThrIleAspIleThrGluLys-----IleSerAsnIleGluGlnLeuLe 1931
Db 6777 TGAACGAGATGCCATAGCAGAAAGAAAGAGATTCAAACTTAGAAGAGCAATAGA 6836
QY 1931 uGlnGlnAlaThrAsnLeuLysGluThrLeuTyrGluArgGluSerLeu-----IleGlnCy 1950
Db 6837 ACAGTTTAGAGAGAACTGGGAAAAATAGAAATGAAAGAGTTCAACAATTATATGCAATTT 6896
QY 1950 slyGluGlnLeuAlaLeuAsnThrGlnIleAspArgGluThrLeuLysSerLysAspLe 1970
Db 6897 AGAAATACAGAAAAAGAAATCTACTACCGCTACAGAACTTGACAGGAAAACAATTT 6956
QY 1970 uAlaLeuGlyLysMetGlu----- 1976
Db 6957 ATTTAAGATGACATGAGAAACTGGGACTTGCATAAAGAAATTCATGATCCATGCTAC 7016
QY 1977 -GlnGluArgAspLysAlaAlaAsnLysValIleAlaLeuThrGluLysMetSerSerLe 1996
Db 7017 TCAAGACCAACATGCTCTATTGGGAAATTTGCTCAAAATATACAGGAAAAAGAGTAGA 7076
QY 1996 uGluGluGlnIleAsnGlnAsnValThrThrLeuLysGluGluGluGlyLysGluThr 2016
Db 7077 AATGACCAATTAATGACAAAGTTACGAAACTCCAG----- 7113
QY 2016 rPheTyrLeuGlnArgProSerLysGlnGlnSerSerSerGlnMetGluLeuArgGln 2036
Db 7114 -----CAGCACTTAAATTAATACACAGATACAAAGGTTATGAGAAAAAATGA 7163
QY 2036 uSerLeuLysThrLysAspLeuGlnLeuGln----- 2046
Db 7164 ACTGATTAAGGATCTTGAACCCCAATAGAAATGTTTGTATGATGATGATCAAGAAATGTGAA 7223
QY 2047 -----GluAlaGluLysGluIleSerGluAlaThrAsnGluIleLysAsnLeuThrAlaLys 2065
Db 7224 GAGAAATAGAGAAAGAAAGAAATAGAGCAATGAAAGTAGTGAAGAAACTTGAACAGGA 7283
QY 2065 sIleSerSerLeuGluGluGluIleLeuGlnAsnAlaSerIleLeuAsnGluAlaValSe 2085
Db 7284 ATTGGCAAAATATGAGCAGAGAAAGACATCAATGAAATGCTCATTCCTCTCAGAA----- 7335
QY 2085 rGluArgGluAsnLeuArgHisSerLysGlnGlnLeuValSerLys-----LeuGln 2102
Db 7336 -GAAGCAGACAGTTTAAACATCAATGATGATGTTATACCTGTAAGAAAGCTGGCTTGA 7394
QY 2102 uGln-----LeuSerLeuThrLeuLysSerAr 2111
Db 7395 ACAGCAAGTAAACCCCTAATGAAAGAAATGACCTTATGAAATATGACTTAAAGAAAC 7454
QY 2111 gaSPHisAlaPheAlaGln-----SerLysArgGluLysAspGln 2124
Db 7455 CAATTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7511
QY 2124 uAlaValAsnLysIleLeuSerLeuAlaGluGln-----IleLysIleLe 2139
Db 7512 AAGTGTGAAAGATTAAGCAATACAGAGAAATAGTGTAAAGCTGCTAATAGTCAATCT 7571
QY 2139 uThrLysGluMetAspLysPhe-----ArgAspSerLysGluSerLeuGln 2154

Db 7572 GAGCAAAAGCAAAACCTGAACTAGAAAGTAGTCTTACAGAGATGCTTAAATCCCTAGA 7631
QY 2154 nGluGlnSerSerHisLeuSerGluGlu-----LeuCyThrTyrLys 2168
Db 7632 AAATGACATATCTTCAATCTTTTGAAGAAATAGCAAGAGTTCATTAATTAATTTGGA 7691
QY 2168 sThrGluLeuGlnMetLeuLysGlnGlnLysGluAspIleAsnAsnLysLeuAlaGluLys 2188
Db 7692 AAACAAGTCTTACAACTTGAAGACACCTGTAGTCCAAAGAGCTTGAAGAACTTACCAATG 7751
QY 2188 sValLysGluVal-----AspGluLeuLeuGln-- 2197
Db 7752 TTATTAACAAATTAAGACATGCAAGAACAGCCAGTTTGAACAGAAATGCTTCAAAA 7811
QY 2198 -----HisLeuSerSerLeuLysGluGln-----LeuAspGln 2208
Db 7812 GAAGATTGTAAACCTTACAGAAATAGTTGAAGAAAAAGTGGCTGCTGCTTGTCTAGTCA 7871
QY 2208 nIleGlnMetGluLeuArgAsnGlu----- 2216
Db 7872 AATCCAACTTGAGCGAGTTCAGAAATATGCAAAATTCGTCAAGATTAATCAAAATTTTC 7931
QY 2217 -----LysLeuArgAsnTyrGluLeuCy 2224
Db 7932 ATCAGAACTGAAAGAACAAATATTCAGAAATTTAAATCACTAAGAGAAATGATGCTGG 7991
QY 2224 sGluLysMetAspIleMetGluLysGluIleSerValLeuArgLeuMetGlnAsnGluPr 2244
Db 7992 GTCAAGATATATCAGCATTAACCTTGAGAAATATCAGAAATAGCAAGCCAGTGTGGAAT 8051
QY 2244 o-----GlnGlnGluGluAspArgValAlaGluArgMetAspIleLe 2258
Db 8052 GCATCTAGTGTGATTTAGAAAAAGAACAGTACAAATTTGCAGAAAAA---AATGTTT 8108
QY 2258 uGlnSerArgAsnGln-----GluIleGlnGluLeuMetGluLysIleSerAlaValTyr 2276
Db 8109 AGAAAAAGAAAGAGCTGCTAGAACTACAGAACTATTTGAGGCAATGAGAAAAAACA 8168
QY 2276 rSerGluGlnHisThrLeuLeuSerSerLeuSerSerLeuGlnLysGluThrGluAl 2296
Db 8169 GAGAGAGAAAGAAAGAAAGAAAGCCCTCAAGATGTTCAAGTTCACAGAACAACTGCA 8228
QY 2296 aHisLysHisCysMetLeuAsnIleLysGluSerLeuSerSerThrLeuSerArgSerPh 2316
Db 8229 GCTATTTAT-----AGCAATGAAAGAGTGTATTTAATGAAT 8270
QY 2316 eGlySerLeuGlnThrGlnHisValLysLeuAsnThrGlnLeuGlnThrLeuLeuAsnLys 2336
Db 8271 CGAGGCTCTTAGAGCTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8330
QY 2336 sPheLysValValTyrArgThrAlaAlaValLysGluAspHis-----SerLeuIleLys 2354
Db 8331 GGCTGAAAACTTCAAGAAAGAGCTTTTGTAAAGAAACAAATATGATCATCTTTCAGAA 8390
QY 2354 sAspTyrGlu-----LysAspLeuAlaAlaGluGlnLys----- 2365
Db 8391 AGACTTAAGCCAAAGTTAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8450
QY 2366 -ArgHisAspGluLeuArgLeuGln-----LeuGlnCysLeu---GluGlnHisGln 2381
Db 8451 AAAGAAAGATGAGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8510
QY 2381 YArgLysTyrSerAspSerAlaSerGluGlu-----LeuLysPheCysGluI 2397
Db 8511 TATATAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8570
QY 2397 eGluPheLeuAsnGluLeuLeuPheLysLysAlaAsnIleIleGlnSerValGlnAspAs 2417
Db 8571 CAATCAAGCTTCACAAAATCTTGTAAATATGAGAAATTAATTAATTAATTAATTAATTAAT 8630
QY 2417 pPheSer-----GluValGlnValPheLeuAsnGlnValGlySerThrLeuGlnGln 2435


```

Db 8831 ACAGACTTATGACAGTGC-----CTTGTGATATATTATACCTCTCTT 3875
Qy 1126 sleuValGluGlyAsnSerGlnAspProIleGluIleAsnAspTyr----- 1141
Db 3876 AAAATGTGAAGTAAATGCGAAGACAAAGAGAATTCGTGTGATTACATTTCTGAAATGA 3935
Qy 1142 -----HisAs 1143
Db 3936 ACATCCAGATTAACAAGATTATAGATAGATTCAGACTTCAGAAATAATATGCACAC 3995
Qy 1143 nleuIleAlaIleuAlaThrGluArgAsnAsnIleMetValCysLeuGluThr----- 1160
Db 3996 TCTTCTCAACAAAGTAAAGAGAAATACAAACAACCTCTGTACTTCAACACGACTAG 4055
Qy 1161 -----GluArgAsnSerIle 1165
Db 4056 CAAGATCTGGGACAGACAGACAGATGATGAAACTTGAATTTGGAGAAAGAAACCTTCC 4115
Qy 1165 uLysGluGluVal-----IleAs 1171
Db 4116 AAAAGAGGAACAGAGTTTTCATCATCTCATCTCAGATGACCAATTTGGAAAGACATTGA 4175
Qy 1171 pLeuAsnThrGlnLeuGlnSerLeuGlnAlaGlnSerIleGluYsSerAspLeuGlnIly 1191
Db 4176 TGTCAATCATTAAGAACAGATTATCTTCTCTGCAAGATCTTGAATAAACTAACTTGAGA 4235
Qy 1191 sProIysGlnAspLeuGlu-----GluGlyGlu 1200
Db 4236 ACAAGTTCAAGAAATTGAAAGGCGCTATATCTTTCGACAGCAAGTTGAAAGAAACTGA 4295
Qy 1200 uValIysLeuLeuLeuGluMetGluLeuLeuYsGlyHisLeu----- 1214
Db 4296 ACAAAACATATGAGGAGAGATCCTACTGTTTACAGAAAGGCTTCAAGCTTATAGTACAGC 4355
Qy 1215 -----ThrAspSerGlnIle 1219
Db 4356 CACGCTCCGCCACAGCTTACTCTGTGATTCGGTGTATATACAGATCTGATGCACAGAG 4415
Qy 1219 uSerIle----- 1221
Db 4416 AACAAATGTAACCTGGAAGTGTGTGAAAAAAGAAATATGATGTACAAATAGACTTTTCTGG 4475
Qy 1222 -----GluYsLeuGlnLeuGlnLeuAsnLeuGlnValAlaThrGluYsLeuGlu 1236
Db 4476 TGAATTGGAGTGAAGAGAGAGAAACAATATCTCTTAAGTTGCTTGAATAACAATACCAAGA 4535
Qy 1236 nThrLeuGlnGluMetLysAsnIleThrIle-----GluArg 1249
Db 4536 ACAATTAGAGAAAGATAGTAAAGTTATGTGTCAATGAGTATAGCATTTGTCACACA 4595
Qy 1249 GAsnGluLeu----- 1252
Db 4596 AACCTAAGCTCTAGAAATATCTGGGGGAAAGAAATACTGCATCATCAACCAAGACACA 4655
Qy 1253 -----GlnThrAspPheGluAspLeuYsAlaGluHisAspSerIle 1266
Db 4656 TGCCTGTGTGCGAAGAACAAACATATTATTATGAATGAATATATATCAGAGATCAAAAT 4715
Qy 1266 uLysGlnAspLeuSerGlnAsnIleGluGlnSerIleGluThrGlnAspGluLeuArgAl 1286
Db 4716 TGGTTTTCAGACTTTTGAACAGCTGAT-----GTGAATTTTAAAGAAAGATTTAAACC 4769
Qy 1286 AlaIleGlnGluLeuArgGluGln-----LysGlnLeuValAspSerPheArgGln----- 1303
Db 4770 ACTTAGTAAGATTAAGAGCAATGAGAAAGAAATTTTATATCAAAATAGATGCCCA 4829
Qy 1304 -----GlnLeuLeuAspCysSerValGlyIleSer-----SerProAs 1316
Db 4830 TGATATACCAAGATCAAAAGAGCTGTGTCTGACTATTTTCAGAGAAATGTTTCCAAAGA 4889
Qy 1316 HisAspAlaValAlaAsnGln-----GluYsValSerLeuGluValAsnSe 1333

```

```

Db 4890 TAAACATTTATAGTAGACAGTATCATGATGATTCAGTACAGTCAAGCATGATGC 4949
Qy 1333 r-----LeuGlnSerGluMetLeuAlaGlyGluArgAspGluLeuGlnTh 1348
Db 4950 TTCTAGACACTAATGTATTAATGAAGACAGTTGGAAGATATAGACAGCAAGCTGTAGC 5009
Qy 1348 rSerCysAlaLeuValSerGluLeuLeuAlaArg-----AlaHisValIysSerVa 1367
Db 5010 ACAATTCGAAGACATCAACAGGACAGCAAGATTTGTTAAAGCAAGCATATATGCGCAAT 5069
Qy 1367 lGluGluGluAsnLeuGluIleThrLysLeuAsnGlyLeuGluYsGluIleLeuGlu 1387
Db 5070 GGAG----- 5073
Qy 1387 yLysSerGluGluSerGluValLeuYsSerMetLeuGluAsnLeuYsGluAspAsnAs 1407
Db 5074 -AGACAGCAGAGAACACAG-----GAACAGCTTCAAGAAAGAGATTAA 5114
Qy 1407 nYsLeuYsGluGlnAlaGluIleGluYsSerSerLysGluAsnGlnPheSerLeuGluGlu 1427
Db 5115 GAGACTTAATAGCAATTAGCCAGAGATCTTCATAGATTAATGAAACCTGTTTACAGA 5174
Qy 1427 uValPheSerGlySerGlnLysLeuValAspGluIleGluValLeuYsAlaGlnLeuY 1447
Db 5175 G-----AGAGAGAGGCTGCTTTAGAGAGCTGGAAGACATTAACAGCTGCTT 5225
Qy 1447 sAlaAlaGluGluArgLeu-----GluIleYsAspArgAspTyrPheGluLeuValGlu 1465
Db 5226 ACCTGAGAGAGAGAGCTGTGTGCTGCGCAACAGCAGCT-----ACGA 5273
Qy 1465 nThrAlaAsnThrAspLeuValGluGluYsLeuGluThrProLeuGlnAlaAspHisGlu 1485
Db 5274 AACACAGATGGAATAAGAAACCAAGAGAGATGGA-----GAACAAACATTTAA 5324
Qy 1485 uGluAspSerIleAspArgSerGluGluMetGluIleYsValLeuGlyGluYsLe 1505
Db 5325 AGAAAGGAATTAAGACAGAAACCTGAGATGCTGCTGCTGATTTGTTCTTAAGAA-- 5382
Qy 1505 uGluYsArgAsnGlnYrLeuLeuGluArgLeuGlnGluGluYsLeuGluLeuSerAsnIly 1525
Db 5383 -----AGTATGCACTCCAGAAACCTAAT-----AATAGACT 5414
Qy 1525 sLeuGluIleLeuGlnLysGluMetGluThrSerValLeuLeuYsAspAsp----- 1542
Db 5415 TTGTAAGATCCCTTGAAGTGTAAAGACAAACAGCAGCTGTGAAGAAACAATTCGTGC 5474
Qy 1543 -----LeuGlnGlnLysLeuGlnSerLeuLeuSerGluAsnIleIleLe 1557
Db 5475 CCATGCTTGGGATTCATGATATGATCTAGTAAAGCCAGTCATCTGCCAGCTTAATTG 5534
Qy 1557 uLysGluAsnIleAspThrThrLeuYs-----HisHisSerAspThrGlnAlaGlu 1574
Db 5535 GAGGTCAGAGAGAGAGCATGTGTAAGATGTCTCATGAGAAACATACAGAGATTAC 5594
Qy 1574 nLeuGlnLys----- 1577
Db 5595 ACATGAATCCATCCCTCTTATTCGGAAGTATATGCCAAGAAATGACATTAAACATGCG 5654
Qy 1578 -----ThrGlnGlnGluLeuGlnLeuAlaLysAsnLeu-----AlaIleAl 1591
Db 5655 GTCAAAAGTAATGAGGAAGAAAGAGCTGTACACAAACACTTGTGAGGAGGTGTTTGC 5714
Qy 1591 AlaSerAspAsnCysProIleThrGlnGlu----- 1601
Db 5715 TGGAACTGAATAGAACCTGAAGAAATGAAAGAACTTATGCTGAACATTAGCTCGACTACA 5774
Qy 1602 -----LysGluThrSerAlaAspCysValHisPr 1611
Db 5775 AACACAGCTTGAAGAAACCTCTGAAGACCTAAGTGAATACAGACTGACCTGAACATGCC 5834
Qy 1611 oLeuGluGluYsIleLeuLeuLeuThrGluGluLeuHisGlnIlyThrAsnGluGlnGlu 1631
Db 5835 GAAAGTGAACAGACAGAGATGATGCGTGAATTCATTAGACAGAAACAAAGAAACAGAGA 5894

```

QY 1631 uTysLeu-----LeuHisGluLysAsnGluLeu---G1 1641
 Db 5895 GTCCCTTAAGTGCACAGAGAACTTCGAGAGCCGCTTCATGAGAGAGTCCAGGGCCAGAG 5954
 QY 1641 uGlnAlaGlnValGluLeuLysCysGluValAlaGlnHisLeuMetLysSerMetIleGlu 1661
 Db 5955 ACACCTAGCTGTGGAGCTCACT---AAGCGTAGGGCGCTCATGTATGCTATGCAATGA 6011
 QY 1661 rLysSerSerLeuGluSerLeuGlnHisGluLysHisAsp-----ThrGluG1 1677
 Db 6012 AAAAAGCTTTTGAAGGCAAAATTCAGGAAAAAAGTATATATAGATGCTGTTCAGCA 6071
 QY 1677 nGlnLeuLeuAlaLeuLysGlnGlnMetGlnValAlaThrGlnGluLysGluLeuG1 1697
 Db 6072 GGAGCTTTGATGTCACAGTAAACAGTTGCAAGAAATGAGGACAGACAGACAGATCCA 6131
 QY 1697 nGlnThrHisGluHisLeuThrAlaGluValAspHisLeuLysGluAsnIleGluLeuG1 1717
 Db 6132 AGAAGAAAGAGAAATTAATCTGTCACAGCAAAAGAGCTATGAAA-----GCAGAGCGAG 6185
 QY 1717 yLeuAsnPhelLysAsnGluAlaGlnGlnLysThrThrLysGluGlnCysIleLeuLeuAsnG1 1737
 Db 6186 C-----CCAGTTGAAACAAACAAATTAATCTACAGGA 6212
 QY 1737 uAsnLysGluLeuGlnGlnSerGlnHisArgLeuGlnCysGluIleGlnGluLeuMetLys 1757
 Db 6213 GACGAGAAATTAATTAAGAGGAAAGAACTAGAGATCAATGTCACAGTGAAGAAAGTACGTA 6272
 QY 1757 sSerLeuLysAspLysGluSerAlaLeuGluThrLeuLysGluSerGlnGlnLysVal11 1777
 Db 6273 TGACCTTCAAAAACAAAGTGAAGCTCTAGAA-----AT 6305
 QY 1777 eAsnLeuAsnGlnGlnMetGluMetValMetLeuGlnGlnMetGluGluLeuLysAsnSerG1 1797
 Db 6306 AGATGTGAAGAACAGTCACTAGT---TTTATACGCTGGAACAGAAAGAAATTAATCTGCA 6362
 QY 1797 nArgThrValIleAlaGlnArgAspGlnLeuGlnAspSerLeuArgGluSerValGluMe 1817
 Db 6363 ACTAATGATTTTAAGACAGCAAAACCAAGCAATGGAAGAAAGTGAAGAAATTAATGAGAAA 6422
 QY 1817 tSerIleGluThrGlnAspAspSerLeuArgLysAlaGlnGlnAlaLeuGlnGlnLysAs 1837
 Db 6423 ATTTTAGATGAGCAAGCCATTCAGACAGAAACATAGAGATGATGATTCACAGCAG--- 6477
 QY 1837 pLysValGlnGluLeuThrSerGlnIleSerValLeu----- 1849
 Db 6478 -CAATACAGAACTAGAACAGCAACTTAAGGTGTCTCCTGATTCAGCCATACAGTGA 6536
 QY 1850 -----GlnGluLysIleSerLeuLeuGlnGlnMetLeuLysAsnValAlaThrVa 1867
 Db 6537 ACATCAAACTAGAGAGGTTCAGACAGTTCAGCAATCATCTGAAAGAAAGAAAGACAGCAATG 6596
 QY 1867 LysGluThrLeuSerGluArgAspAspSerLeuAsnGlnSerLysGlnHisLeuPheSerG1 1887
 Db 6597 CAGTGTGCTTTTGTCTCTTAAGAGAGAGCTTCAGAGGGATATACAGAAAGAAAGATGAAGA 6656
 QY 1887 uIleGluThrLeuSerLeuSerLeuLysGluLysGluPheAlaLeu----- 1902
 Db 6657 AATAGAGAAACGTGAGATTCAGAGTAAGAGAACTGAGCGCTCTTGTAGAGAGCGC 6716
 QY 1903 -----GluGlnAlaGluLysAspLysAlaAspAl 1912
 Db 6717 AAAACACTTTCAGAGCTGTAGAGCTTAACCAAGAAATTCCTCAGAGATCAATTCAGAGC 6776
 QY 1912 AlaAlaGlyThrIleAspIleThrGluLys---IleSerAsnIleGlnGluLeu 1931
 Db 6777 TGAAGAGAGTGCATAGACAGAAAGAAAGAGATTAACAATCTAGAGAGCAATTAAGA 6836
 QY 1931 uGlnGlnAlaThrAsnLeuLysGluThrLeuLysGluArgGluSerLeu---IleGlnCys 1950
 Db 6837 ACAGTTTAGAGAGAACTGGAAGAAATTAAGATGAAGAGTTCAACAATTCATATGCAAT 6896

QY 1950 sLysGlnGlnLeuAlaLeuAsnThrGlnHisLeuArgGluThrLeuLysSerLysAspLe 1970
 Db 6897 AGAAATACAGAAAGAAAGATTCATACCCGCTACAGAAACTTGAACAGAAACAAAT 6956
 QY 1970 uAlaLeuGlyLysMetGlu----- 1976
 Db 6957 ATTTAAGATTCATGAGAGAAACCTGGAGCTTCATTAAGAAATCTGATCCATCTTAC 7016
 QY 1977 -GlnGluArgAspGlnAlaAlaAsnLysValIleAlaLeuThrGlnLysMetSerSerLe 1996
 Db 7017 TCAGAGCCACATGCTGCTCTTTGGGAAATTCCTCAATATATACAGAAAGAGAGTGA 7076
 QY 1996 uGlnGlnGlnIleAsnGlnLysAsnValThrThrLeuLysGluGluGluGluLysGluTh 2016
 Db 7077 AATTGACCAATTAATGAACAAAGTATGAGAACTCCAG----- 7113
 QY 2016 rPheLysThrLeuArgProSerLysGlnGlnSerSerSerGlnMetGluLeuArgG1 2036
 Db 7114 -----CACCACTTAAATTTCAACACAGATTAACAAGTTATGAAAGAAATTAATCA 7163
 QY 2036 uSerLeuLysThrLysAspLeuGlnGlu----- 2046
 Db 7164 ACTGATTAAGGAGCTTGAAGAACCAATAGAAATGTTGATGATGATCAAGAAATGTCGAA 7223
 QY 2047 -----GluAlaGluLysGluIleSerGlnAlaThrAsnGluIleLysAsnLeuThrAlaLys 2065
 Db 7224 GAGAAATAGAGAAAGAAATTAAGAGACCTCAATGAGAGTGAAGTGAAGAAATTCACAGCA 7283
 QY 2065 sLysSerSerLeuGlnGluGluThrLeuGlnAsnAlaSerIleLeuAsnGlnAlaValSer 2085
 Db 7284 ATTTGCAAAATTTTGAGACAGAAACATCAATGATGCTCATCTCCCTCAGAA----- 7335
 QY 2085 rGluArgGluAsnLeuArgHisSerLysGlnGlnLeuValSerGlu-----LeuG1 2102
 Db 7336 -GAAGCAGACAGTTTAACATCATGATGATGATTAAGCGAAAGCTGCTTGA 7394
 QY 2102 uGln-----LeuSerLeuThrLeuLysSerAr 2111
 Db 7395 ACACCAAGTGAACCCCTAATGAGAAATGACCTTCATGAAGAAATGCTTAAAGAAAC 7454
 QY 2111 gAspHisAlaPheAlaGln-----SerLysArgGluLysAspG1 2124
 Db 7455 CAATTTAAATTAAGATCAATCAACAGCAATTAATTCAGCTTAAAGAGAAAGT---GA 7511
 QY 2124 uAlaValAsnLysIleLeuSerLeuAlaGlnGlu-----IleLysIleLe 2139
 Db 7512 AAGTGTGAAGAAATTAAGATCAATCAACAGCAATTAATGATGATGCTTATATATCATCT 7571
 QY 2139 uThrLysGluMetAspGluPhe-----ArgAspSerLysGluSerLeuG1 2154
 Db 7572 GAGCAAGACAAACCTGAAACAGAAAGTACCTTACAGAGAGATGCTTAAATCCCTAGA 7631
 QY 2154 nGlnGlnSerSerHisLeuSerGluGlu-----LeuLysThrLysLys 2168
 Db 7632 AAATTCAGCAATCTCAATCTTTTGAAGAAATGCAAGAGTTCCTCAATTAATTTGA 7691
 QY 2168 sThrGlnLeuGlnMetLeuLysGlnGlnLysGluAspIleAsnAsnLysLeuAlaGluLys 2188
 Db 7692 AACCAAGTTGCTCAACTTGAAGACACTGTTAGTCAAGAGACTTAACCTTACCCAGTG 7751
 QY 2188 sValLysGluVal-----AspGluLeuLeuGln--- 2197
 Db 7752 TTATTAACAAATTAAGAAAGATGCAAGAAAGCCAGCTTTCGAAACAGAAATGCTTCAAAA 7811
 QY 2198 -----HisLeuSerSerLeuLysGluGln-----LeuAspG1 2208
 Db 7812 GAAGATTTGAACCTTACAGAAAGAAATGTTGAAGAAAGAGTGTCTGCTTTCAGTCA 7871
 QY 2208 nIleGlnMetGluLeuArgAsnGlu----- 2216
 Db 7872 AATTCACCTTGAAGAGCTTCAAGAAATATGCAAAATTCGTCAAGATATACAAATTC 7931
 QY 2217 -----LysLeuArgAsnLysGluLeuGly 2224

Db 909 AGAGTTG-----ACAGAACAGTCAAAAATTACAGATTCAT 947
 Oy 421 gatValThrTrpAlaProGlyLysIleGlnAsnSerLeuHisAlaSerGlyValSerAs 441
 Db 948 TCAGCAATTACAGAGCTTAGTGAACCTGTGAGAACACAGCTACTATGAGCAGACCTCAGA 1007
 Oy 441 p-----PheAspMetLeuSerAr 447
 Db 1008 CTACTACAAAGCCAAACACAGATCTCTCACTCATCAACAGCAGCTTGAGAACAAACACCA 1067
 Oy 447 gleupProGlyAsnPheSerLys----- 454
 Db 1068 CTATTAGAGAGATTATCAACAAAAGAGAACAGCTTCACATGCAATTAAGTTCTTGCA 1127
 Oy 455 ----LysAlaLysPheSerAspMetProSerPheProGluIleAspAspSerValLysThr 473
 Db 1128 ACAGAAAATTAAGTATATGAATGACACAGATTAATAAAAGTGAAGAACTCAATTAAGA 1187
 Oy 473 rglupSerAspPheAspAlaLeuSerMetLeuAspSerAsnGlyIleAspAlaG 493
 Db 1188 AGAAATACAGAAAAGACACATCATGAGAAATTAACACAAAATATATAGAACAGA 1247
 Oy 493 u-----TrpAsnLeuAlaSerLysValThrHisArgGluLysThr-----SerIe 508
 Db 1248 AAGAAACTCTTGACTTAAGGATTAATTAACACTGCTGATTAATTAAGAGAAAT 1307
 Oy 508 uHisGlnSerMetIleAspPheGlyGln-----IleSerAs 520
 Db 1308 ACAGAACAGATTGTGCACAAAGACACCAAGAAATTAATAACATGAATTTAGAGCTGACTTA 1367
 Oy 520 pSerValGlnPheHisAspSerSerLysGlnAsnGlnLeuGlnIleuProLysAspSe 540
 Db 1368 TTCTAAGCAAAAAGAAAGACAGCTCTGAGAAATTAACAGTTAATG----- 1416
 Oy 540 rglYaspMetIaGluCysArgLys-----AlaSerPheGluLysGluI 555
 Db 1417 -GGGACAGTCGAGAACCTTCAGAGAAATATCAATAAGACAGCCAGCTGGAACTATAT 1475
 Oy 555 eThrSer---LeuGlnGlnIleuGlnSerLys---GluGluGluLysArgLysGluLeuVal 573
 Db 1476 AGTACACCAATGGAACAAAGAAACACAAAGAAAGTTAGAACAACTCCGGCAGAGCTGCA 1535
 Oy 573 IeGlnSerPheGluLeuLysIleAlaGluLeuGlnGlnIleu-----SerValLysAl 591
 Db 1536 TGAGATGATGTGGACAGACATAGTGCATGAATGAACAAAGATTAATGAACACACATGCG 1595
 Oy 591 aLysAsnLeuGluMetValThrAsnSerArg-----GluHisSerIleAsnAlaG 608
 Db 1596 ACAGATGAGAGAAATGAACACGGGCATAGGAGAAATGAGAAATGCTTTAAGGTCAATA 1655
 Oy 608 uValGlnThrAspValGluLysGluValAlaArgLysGluMetSerValLeuGluLysAspSe 628
 Db 1656 TTCAATATATTACAGTTAATGAAGATCAATTAAG-----TTAATGATGT 1700
 Oy 628 rglYTrAsnAlaSerAsnSerAspLeuGlnAspSerSerValAspGlyLysArgLysSe 648
 Db 1701 GGCAATTAATGAATGAATGAATTAATTAATGCAAGTACTACTCACTCAAAAGAAACCTCAA 1760
 Oy 648 rSerSerHisAspGluCysIleGluHisArgLysMetLeuGluGlnLysIleValAspLe 668
 Db 1761 GGAAGAACTAGACATTAATTTTAGAGAAAGTGTGCTTACAGAGACCTTGAAACCT 1820
 Oy 668 uGluGluPheIleGluAsnLeuAsnLysSerGluAsnAspLysGluLysSerSerG 688
 Db 1821 TGTGAGAAATGAGCTTTTCAAGGAGAACAGATTCAGAGAGCTAGACACATTAAGCTCA 1880
 Oy 688 uGlnAspPheMet-----GluSerIleGlnLeuGluLysGluAlaIleMetAl 703
 Db 1881 ACAAGAAAGTAAGCTTAATGAAGACATTAAGTCCCTTAGTACAGTGAAGATTTGAAGC 1940
 Oy 703 agLu----- 704
 Db 1941 TAGAATTTGTTCTGCACTGTAATCCAGAAAGAACTAGAAATTAATAACATGAAGCAGAGT 2000

Oy 705 -----LysAlaAsnAlaLeu----- 709
 Db 2001 TACAAATTACAAATAAACTGTAATGTTAGAAAAAGAAAGATGCTGTGTAGACAG 2060
 Oy 709 ----- 709
 Db 2061 AATGGCTGAATCACAAGAAAGCTGAATTAGAGAGCTGAGAACACAGCTTATTATGTC 2120
 Oy 710 ----GluGluLeuAlaLeuMetArgAspAsnPheAsp-----AsnIleI 723
 Db 2121 CGAAGAGAGCTTCCAAACCTGAGAGAGATTTAGAAATTAAGAACATCGAATTAATTTGA 2180
 Oy 723 eleuGluAsnGluThrLeu-----LysArgGluIleAlaAspLeuLysSe 739
 Db 2181 AAAACTTAAGATTAATTTAGCGATTCATTAACACAGCATGATGCTTACAGATCA 2240
 Oy 739 rLeuLysGluAsnGlnGluThrAsnGluPheGluIleLeuGluLysGluThrGlnLysG 759
 Db 2241 AATGAGTCAAAAGATAGAAACATGACGATT-----GAAAGGACAAATTTGATTAAC 2291
 Oy 759 uHisGluAlaGlnLeuIleHisGluIleGlySerLeuLysLysLeuValGluAsnAlaG 779
 Db 2292 TAAAGCAATCAATTAATTTGGAATTTCCAAAGCTTAAGATTTACAGCAGCT----- 2346
 Oy 779 uMetYTrAsnGluAsnLeuGluGlu-----AspLeuGluThrLysThr 793
 Db 2347 -CTTGTAATTTCAAAGTCAAGAAATGACTTCAATCAATGAACTTCAAAAGAAAT 2405
 Oy 793 rLysLeuLeuLysGlu-----GlnGluI 801
 Db 2406 TGAATATCTCAGACAAAGAAAGAAAGGCTGACCTGAGACAAAGATTCAGAAAT 2465
 Oy 801 eGlnLeu-----AlaGluLeuArgLysArgAlaAspAsnLeuLys 815
 Db 2466 ACACCTTAACAGAAATTTGTAGAAAAACAGATGAAGAAAGAAAGATGATTTCAAGC 2525
 Oy 815 sLysValArgAsnPheAspLeuSerValSer---MetGlyAspSerGluLysLeuGlySe 834
 Db 2526 AAAATTTGCCAATCTGAAACAGACAGATAGCATTTCTTAAGATGAAGAAACCCCTGA 2585
 Oy 834 uGluIlePheGlnLeuLysGlnSerLeuSerAspAlaGluAlaVal----- 849
 Db 2586 AGACATGTGAAATATCAATCTCTGTAGCCAAAGAAAGAAATGATTTCTTAGACTC 2645
 Oy 850 -----ThrArgAspAla-----GlnLysGluCysSerPheLeuArgSerG 863
 Db 2646 CATTAACTCCAAATCCAAAGACTCTGTGTGGAAAAAGAAATGAAATCTTTATGAGCA 2705
 Oy 863 uAsn-----LeuGluLeuLysGluLysMetGluAspThrSerAs 876
 Db 2706 AATGAGAGACCTCAAAACAAATGATTTACAGCTAATTAAGAAAGATGAAAGCAAGAAAGAA 2765
 Oy 876 nTrpYTrAsnGlnLysGluLysAlaAlaSerLeuPheGluLysGlnLeuGluThrGluLys 896
 Db 2766 CACTTTTTCATTTGCGAAAAAACTTTGAAGTTACTATCAAGAGCTTACAAAGAGAGTA 2825
 Oy 896 sSerAsnYTrLysLysMetGluAlaAspLeuGln-----LysGluLeuG 911
 Db 2826 TGCTTCCTTTCACAAAGTAAAGATGATTTAGAGAACAGTAAATAATTAACAGGAATTAAG 2885
 Oy 911 n-----SerAlaPheAsnGluIleAsnYTrLeuAsn----- 921
 Db 2886 GTATTAAGAAACTTAAGCACTTATATGAAGAGCTTCATTTCCAAAGAAATTAATCCAAC 2945
 Oy 921 ----- 921
 Db 2946 TACAGTGAATGAAGATTTCTGTGATGAGAGCAAAACTTTGTAGACAGAAACATT 3005
 Oy 922 ----GlyLeuLeuAlaGlyLysValProArgAspLeu----- 933
 Db 3006 GGAATGAGTGAAGTGTGTAAGAAAGATACACAGAACTGATGAAAAAACTTGAGGTAAC 3065

QY 934 -----SerArgValGluLeuGluLysValLysGluPheSerLysGluLeuGluLys 951
 Db 3066 CAAAGCAGAGAAATTAAGCTGTCACAGACAGCTGTCGATCTTGAACAAATGGAACA 3125
 QY 951 salaleuGluLysAsnAlaLeuGluAsnGluValThrCysLeuSerLysLys 970
 Db 3126 GAAACATGCTGAGATTTGTTTCTAAATGGAAGAGTTAACTTTAAAGCAAGAGAAAGA 3185
 QY 971 -----PheLeuProAsnGluValGluCysLeuLysAsn 981
 Db 3186 ACAAGTTTCATTGAGATGAGTAGAGCTAGAAATCATATTAAACCAACAAGGCGAGAAAA 3245
 QY 982 -----GlnIleSerLysAlaSerGluLysIleMetLeuLeuLysGlu 995
 Db 3246 TGACAGTCATGATACTACTCAAGTAACTCTTTATTAGATGAGAGTTGTGACCATGACAG 3305
 QY 995 ngLugly-----GluHisSerAlaSerIleIleSerLys 1006
 Db 3306 CAGGGTCTCTGAGAGATCAGTTCTTAAAGTAAATAAAGTTTGAGAGATCAAAAT 3365
 QY 1007 -----GlnGluIleIleMetGlnLysGlnLysGln 1017
 Db 3366 AATGTGGAAGATAAATGTTCTTTGAAATATGCTGTGAGAGAAATGAAAGCAAGA 3425
 QY 1018 -----IleLeuGlnLeuThrAspGluValThrHisThrGlnSer-----LysValGlu 1033
 Db 3426 ACAGTTGATTTGATGACTTACCATCTGTACACAAAGATCATCTAGAGCACTCA 3485
 QY 1033 ngLinhThrGluGlnLysLysLeuGlu-----MetLysLysMetHisAspAs 1048
 Db 3486 ACCAAGTGAATGATTAACCTTCAGAAAGAACTCAATGATCTTAAATCAGAAAGAAATGA 3545
 QY 1048 pleupheGluLysLysLysLysSerGluValGlu 1061
 Db 3546 TTTA-----AGGCTACAGATGAGAGCCACCAAGCCATTGCTCTCTCTGT 3590
 QY 1062 -----AspLeuLeuArgGlu-----MetGluAsnLeuLysGlyThr 1073
 Db 3591 TTATTCACTCATGTCGATCAGAGTTGCTGCAATATATGAAATGAAAGATTAAGCTCT 3650
 QY 1074 -----MetGluSerValGluValLysIleAlaAspThr 1084
 Db 3651 TTGCAGCTTTAAAGAGAGCTTATTTTCTCAGAGAGAAAGATCAAGAACTTCAGAA 3710
 QY 1084 LysHisGluLeuGlu-----GluThrIleArg-----AspLysGluGlnLe 1098
 Db 3711 AATACACAGTTAGACTGACACTATGAAACACAAAGAAACAGGTGAGAGAAAGCC 3770
 QY 1098 uLeuHis-----GluLysLysLysPhePhe 1106
 Db 3771 TTTTACATCTGCTCATTTGAAAACTTCAAAAGCAGTGTGAAAGATTTCTTATTTT 3830
 QY 1106 eGlnAlaMetGlnThrIlePheProIleThrProLeuSerAspSerLeuProPheSerLys 1126
 Db 3831 ACAGACTTTTTCAGAGTCTC-----CTTGGTGAATTTATATCTCTCTCTT 3875
 QY 1126 sleuValGluLysAsnSerGlnAspProIleGluIleAsnAspLys 1141
 Db 3876 AAAATGGAAGTAAATGCAACAAAGCAAGAAATTCGTGATTTACATTTCTGAAATGA 3935
 QY 1142 -----HisAs 1143
 Db 3936 AGATCCAGAAATTAACAAGATTATAGATTAAGTCAAGACTTCAACGAAATATGCAAC 3995
 QY 1143 nLeuIleAlaLeuAlaThrGluArgAsnAsnIleMetValCysLeuGluThr 1160
 Db 3996 TCTTCTCAACAAAGTAAACAGAGAAATACAAACAACTCTGTTCTCAACAGCACTAAG 4055
 QY 1161 -----GluArgAsnSerLys 1165
 Db 4056 CAAGATCTGGGACAGACAGACAGATGATGAACCTTGAATTTGAGAGAAAGAAACCTTCC 4115
 QY 1165 uLysGluGlnVal-----IleAs 1171

Db 4116 AAAAGAGAAACAGAGCTTTTATCAATCCATCTCAGATGACCAATTTGGAAGACATGA 4175
 QY 1171 pleuAsnThrGlnLeuGlnSerLeuGlnAlaGlnSerIleGluLysSerAspLeuGlnLys 1191
 Db 4176 TGTCAATCATTAAMACAGAGTATCTTCTGCAAGATCTGAAATACTAACTTGAGAGA 4235
 QY 1191 sProLysGlnAspLeuGlu-----GluGlyGlu 1200
 Db 4236 ACAAGTTCAAGATTAAGAAAGCCTCATATCTCTTTCAGCAGCAACAGTTGAAGAACTGA 4295
 QY 1200 uValLysLeuLeuLeuGluMetGlnLeuLeuLysGlyHisLeu 1214
 Db 4296 ACAAAATATGAGCAGAGATTCACCTGTTTACAGAAAGAGCTTCAAGCTTGTAGATC 4355
 QY 1215 -----ThrAspSerGlnLe 1219
 Db 4356 CACGCTCCGCAAGCTTACCTGTTGATTCGGTGTATTTACAGAAATGATGACAGACAG 4415
 QY 1219 uSerIle 1221
 Db 4416 ACAATGTACCTGGAAGTGTGTGAAAAAGATATATGTTGATCAATAGAGTTTCTGG 4475
 QY 1222 -----GlnLysLeuGlnLeuGluAsnLeuGluValThrGluLysLeuGlu 1236
 Db 4476 TGAATTTGGAGTGAAGAGAGAAACAAATATCTTAAGTTGCTTGAAGAAACATACCAAGA 4535
 QY 1236 nThrLeuGlnGluGluMetLysAsnIleThrIle-----GluArg 1249
 Db 4536 ACAATTAGAAAGAAAGATGCTAAGTTATGTGCAATGAGTATAGCATTTGCTCAACA 4595
 QY 1249 gaasGluLeu 1252
 Db 4596 AACTGAACTCTCTAGATATCTGGGAAAAAAGAAATCTGCATCATCAACCAAGCACA 4655
 QY 1253 -----GlnThrAsnPheGluAspLeuLysAlaGluHisAspSerLys 1266
 Db 4656 TGCTGTGTGCACAGAAACAAACATATTTATGAAATGAATATATATACAGAGATCAAT 4715
 QY 1266 uLysGlnAspLeuSerGluAsnIleGluGlnSerIleGluThrGlnAspGlnLeuArgAl 1286
 Db 4716 TGGTTTCAGACTTTTGACAGACAGTGTAT-----GTGAATTTAAAGAAATTTTAAAC 4769
 QY 1286 alaGlnGlnGluLeuArgGluGln-----LysGlnLeuValAspSerPheArgGln 1303
 Db 4770 ACTTAGTAAGAGTTAGAGAAACATGGAAGAAATTTTATCAATATAGTATCCCA 4829
 QY 1304 -----GlnLeuLeuAspCysSerValGlyIleSer-----SerProAs 1316
 Db 4830 TGATATACCAAGATCAAAAGCAGCTGTGCTGACTATTTTCAAGAAATGTTCTCAAAAGA 4889
 QY 1316 nHisAspAlaValAlaAsnGln-----GluLysValSerLeuGluValAlaAsnS 1333
 Db 4890 TAAACATTTATTAAGTTAGACAGCTATTCATGATGATGATTTTCACTGTCAGCATGATGC 4949
 QY 1333 r-----LeuGlnSerGluMetLeuArgGlyLysArgAspGluLeuGlnThr 1348
 Db 4950 TTCTAGACAACTAATGTTGATGAAAGAAAGCTTGAAGATATGAGACAGCACTTGTAGC 5009
 QY 1348 rSerCysLysAlaLeuValSerGluLeuGluLeuArg-----AlaHisValLysSerVa 1367
 Db 5010 ACAATACCAAGAAATCAACAGCAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 5069
 QY 1367 IGluGluGluAsnLeuGluIleThrLysLysLeuAsnGlyLeuGluLysGluIleLeuGlu 1387
 Db 5070 GAG----- 5073
 QY 1387 LysSerGluGlnSerGluValLeuLysSerMetLeuGlnLysLeuLysGluAspAsn 1407
 Db 5074 -AGACAGCAGAGACAGCAG-----GAAACGCTACACAGAGAGATTA 5114
 QY 1407 nLysLeuLysGluGlnAlaGluLysSerSerLysGluAsnGlnPheSerLeuGluGlu 1427

Db 5115 GAGACTTAATAGACATTAGCCAGAGATCCTCATAGATATGAAAACTGGTTTCAGA 5174
 Qy 1427 uValPheSerGlySerGlnLysLeuValAspGluIleGluValLeuLysAlaIleLeu 1447
 Db 5175 G-----AGAGAGAGGGCTTTTAGAGAGCTGGAAGCATTAAAGCAGCTGCTCTT 5225
 Qy 1447 sAlaIleGluGluArgLeu-----GluIleLysAspArgAspLysPheGluLeuValG 1465
 Db 5226 AGCTGGAGAGAGAGAAAGCTGTGTGTGAGCTGGCAACAGCACT-----ACGCA 5273
 Qy 1465 nThrAlaAsnThrAsnLeuValGluGlyLysLeuGluThrProLeuGlnAlaAspHisG 1485
 Db 5274 AACACAGATGGAATGAAACCAAGAGAAAGTTGAA-----GAACAAACATTATTA 5324
 Qy 1485 uGluAspSerIleAspArgAspSerGluMetGluIleLysValLeuGlyGluLeu 1505
 Db 5325 AGAAAGAAATTTAGACAAACAACTGAAAGATGTGCTCTGAGATTTGTCTAATGAA-- 5382
 Qy 1505 uGluArgAsnGlnTyLeuLeuGluArgLeuGlnGluLysLeuGluLeuSerAsnLys 1525
 Db 5383 -----AGTATGCACTCCAGAAAGCTAAT-----AATAGACT 5414
 Qy 1525 sLeuGluIleLeuGlnLysGluMetGluThrSerValLeuLeuLysAspAsp----- 1542
 Db 5415 TTTGAGATCCTCTTGAAGAGTTGTAAAGCAACACAGCAGCTGTTGAAGAACAAATGGTCG 5474
 Qy 1543 -----LeuGlnGlnLysLeuGlnLysLeuSerLeuSerGluAsnIleIle 1557
 Db 5475 CCATGCTCCTGGGATTCATAGATAGATCTAGTAAAGCCAGTCATCTGCGACGCTAATTTG 5534
 Qy 1557 uLysGluAsnIleAspThrThrLeuLys-----HisHisSerAspThrGlnAlaG 1574
 Db 5535 GAGGTCAAGAACAGAGGATCTGTAAACTCATGTCTCATGAGCAACATACAGAGTTAC 5594
 Qy 1574 nLeuGlnLys----- 1577
 Db 5595 AGATGAATCCATTCCTCTTATTCTGAAAGTGAATGATGCCAAGAAATGACATTAAACATG 5654
 Qy 1578 -----ThrgInGlnGlnLysLeuAlaLysAsnLeu-----AlaIleAl 1591
 Db 5655 GTCAAAAGTAACTGAGAGAGAGAACAGAGCTGTACACAGACTTGTGAGAGAGGTTTGC 5714
 Qy 1591 aAlaSerAspAsnCysProIleThrGlnLys----- 1601
 Db 5715 TGAACATGAATAGACCTGAAATAAGAACTTATGCTGAACATTAGCTCTGACATCA 5774
 Qy 1602 -----LysGluThrSerAlaAspCysValHisPr 1611
 Db 5775 AGCAGCAGTGAATAAACTCTAGAAAGCCATAGTAAAGTAACTGACAGCTTGAACATGC 5834
 Qy 1611 oLeuGluGlnLysIleLeuLeuLeuThrGluLeuHisGlnLysThrAsnGluGlnG 1631
 Db 5835 GAAAGTGAACAGACAGAGTTGATGCTGAGTCATTAGACAGAAACAAAGAACACAGA 5894
 Qy 1631 uLysLeu-----LeuHisGluLysAsnGluLeu--G 1641
 Db 5895 GTCCCTTAAGTGCACAGAGAACTTCGAGAGCGCCTCATGAGAGTCCAGGCGCAAGA 5954
 Qy 1641 uGlnAlaGlnValGluLeuLysCysGluValGluHisLeuMetLysSerMetIleGlu 1661
 Db 5955 ACACCTAGCTGTGAGCTCACT---AAGGCTGAGGGCGTCAATGAGCTATGACAGAGA 6011
 Qy 1661 rLysSerSerLeuGlnLysLeuGlnHisGluLysHisAsp-----ThrgLysG 1677
 Db 6012 AAAAATCTTTTAAAGGCAAAATTCAGAGAAACAACTGATTAATAGATCGCTTGACCA 6071
 Qy 1677 nGlnLeuLeuAlaLeuLysGlnGlnMetGlnValAlaThrGlnGlnLysGlnLysG 1697
 Db 6072 GGAGTTGTATGTGCAAGTACAGAGTTCAGAAATTTGAGAGGAGCAACAGCAATGCA 6131
 Qy 1697 nGlnThrHisGlnLysLeuThrAlaGluValAlaAspHisLeuLysGlnLysGlnLysG 1717
 Db 6132 AGAAGAAAGAGAAATTAAGTCTCAGACAAAGAGAGCTATGAAA-----GCAGAGGCGAG 6185

Qy 1717 yLeuAsnPheLysAsnGluAlaGlnGlnLysThrThrLysGluGlnCysLeuLeuAsnG 1737
 Db 6186 C-----CCAGTTGAACACAACTTACTACAGCA 6212
 Qy 1737 uAsnLysGluLeuGlnLysSerGlnHisArgLeuGlnCysGluIleGluLeuMetLys 1757
 Db 6213 GACAGAAAAATTAATGAAGAAAAAACTAGAACTACAACTCAAGCTGAAAAAATACGTA 6272
 Qy 1757 sSerLeuLysAspLysGlnLysLeuGlnLysLeuLysGlnLysValI 1777
 Db 6273 TCACCTTCAAAAACAAAGTGAAGCTCTAGAA-----AT 6305
 Qy 1777 eAsnLeuAsnGlnGlnMetGluMetValMetLeuGlnGlnLysLeuLysAsnSerL 1797
 Db 6306 ACATGTGGAAGAACAGTCAAGTACAG--TTATAGAGCTGGAACAGAAAAAATACGTA 6362
 Qy 1797 nArgThrValIleAlaGluArgAspGlnLeuGlnAspLeuArgGluSerValGlu 1817
 Db 6363 ACTAATG----- 6369
 Qy 1817 tSerIleGluThrGlnAspAspLeuArgLysAlaGlnGlnAlaLeuGlnGlnLysAs 1837
 Db 6370 -----GATTTAAGACAGCAAAACCAAGCATTGGAAGAACAGTTAGA 6410
 Qy 1837 pLysValGlnGlnLysLeuThrSerGlnLysSer-----ValLe 1849
 Db 6411 AAAAAATGACAAAAATTTTATAGATGAGCAAGCCATTGACAGAGAACATGACAGATGAT 6470
 Qy 1849 uGlnGlnLysLysSerLeuLeuGlnLysAsnGlnMetLeuTyraAsnValAlaThrValLysG 1869
 Db 6471 CCACAGGAATATACAGAAACTGAAACAGCACTT-----AAGTTGTTCCCTGATTCGA 6524
 Qy 1869 uThrLeuSerGluArgAspAspLeuAsnGlnSerLysGlnHisLeuPheSerGlnLysG 1889
 Db 6525 GCCTATCAGTGA-----CATCAAACTGAGAGAGTTGA 6557
 Qy 1889 uThrLeuSerLeuSerLeuLysGluLysGlu-----PheAlaLe 1902
 Db 6558 ACAGTTAGCAATTCATCTCAAAAGAAAAAACACAAATTCAGTACGTTTCTCTCTAA 6617
 Qy 1902 uGlnGlnAlaGluLysAspLysAlaAspAlaAlaArgLysThrIleAspIleThrGluLys 1922
 Db 6618 AGAGCAGCTTCAAGAGATATCAAGAAAGAAATGAGAAATGAGAAATGAGAGTTGCA 6677
 Qy 1922 sIleSerAsnIleGluGlnGlnLeuGlnGlnAlaThrAsnLeuLysGluThrLeuTy 1942
 Db 6678 ACTAAGAGAACTGAGACAGCGCCTTCTGTGAGTGCAGATACTTTCAAAAGTAGAGCA 6737
 Qy 1942 rGluArgGlu---SerLeuIleGlnCysLysGlnGlnLeuAlaLeuAsnThr-----G 1959
 Db 6738 CCGAAAAACCTTGGAGCTGTGAAGCTTAAGACCAATTTGCTCCTAGAAGTCAATTCGA 6797
 Qy 1959 uHisLeuArgGluThrLeuLysSerLysAspLeuAlaLeuGlnLysMetGluGlnGlu-- 1978
 Db 6798 GGCTGAACAGATGCCATAGACAGAAAGAAAGAAATGAAACTTACAAACTTGAAGACAAT 6857
 Qy 1979 -----ArgAspGluAlaAlaAsnLysValIleAlaLeuThrGlnLysMetSerSe 1995
 Db 6858 AGAACAGTTTGAAGAACTGGAATAATGAAGATGAAGAAAGTTCAACATTAATATGCA 6917
 Qy 1995 rLeuGlnGlnGlnIleAsnGlnAsnValThrThrLeuLysGlnGlnLysGlnLysG 2015
 Db 6918 ATTACAAATACGAAAAAGAAATCTACTACCCGCTTACAGAACTTGAACAGAAAAACAA 6977
 Qy 2015 uThrPheTyrrLeuGlnArgProSerLysGlnGlnSerSerGlnMetGlnGlnLeuAr 2035
 Db 6978 ATTATTT-----AAGATGACATGAGAAACTGGG 7007
 Qy 2035 gLysSerLeuLysThrLysAspLeuGlnGlnGlnLysGlnLysGlnLysGlnLysG 2049
 Db 7008 ACTTGCCATTAAGGAATCTGATGATGATCTACTCAAGCAACAACTGTGCTATTTGGAA 7067

QY 2050 -----LysGluIleSerGluIaThrAsnGluIleLysAsnIleuThAlaLysIleSe 2067
Db 7068 ATTTGCTCAATATACAGAGAAAAGAGTAGAAATTCACCAATTAATGACAGAGTAC 7127
QY 2067 rSerLeuGluGluGluIleLeuGlnAsnAlaSerIleLeuAsnGluAlaLysSerGluAr 2087
Db 7128 GAAATCTCACAGCAACTT-----AAATTCACAGATACAGATGATTATTCAGAA-- 7179
QY 2087 gGluAsnLeuArGHisSerLysGlnGlnLeuValSerGluLeuGlnIleuSerLeuTh 2107
Db 7180 -----AAAAATGACATGATACAGGATCTTGAAACCCAAATAGAAATG 7220
QY 2107 rLeuLysSerArGAspHisAlaPheAlaGlnSerLysArGLeuLysAspGluAlaValAs 2127
Db 7221 TTTGATAGTATCAGAAATGCTGTGAAGAGAAATGAGAAAGAAATATGACAGCTCAA 7280
QY 2127 nLysIleAla----- 2130
Db 7281 TGAAGTATTCAGAAACCTTCACAGAGATTCGCAATTTTGACAGAGAACATCAATGAA 7340
QY 2131 -----SerLeuAlaGluGluIleLysIleLeuThrLysGluMetAspGluPheArGAs 2148
Db 7341 TGCTCATTCCTCTCAGAGAGACAGACAGATTAAACATCAATGGATGGTTATAGC 7400
QY 2148 pSerLysGluSerLeuGlnGlnIleuSerHisLeuSerGluGluMetCysThrTyLys 2168
Db 7401 TGAAGAGCTGGCTTGGAACAGCAAGTAGAACCCTATATAGAAATGACCTTCATGAA 7460
QY 2168 sThr-----GluLeuGlnMetLe 2174
Db 7461 AAATGACTTAAGAAACCAATTTAAATGATCAGCTAACAGAGAAATTAATTCAGCTT 7520
QY 2174 uLysGlnIleLysGluAspIleAsnAsnLysLeuAlaGluLysValLysGluValAspG 2194
Db 7521 AAAGAGAGAGAGCTGAAAGCTG-----GAAAGATTCAGAACCATACAGAA 7555
QY 2194 u-----LeuLeuGlnHisLeuSerSerLeuLysGluIleuAspGlnI 2209
Db 7556 GAATAGTGTAACTGAGCTATATGATCATCTGAGAAACAGCAACCTGATAGAAAGTGT 7625
QY 2209 eGlnMetGlu-----LeuArGAsnGluLys-----LeuArGAsnTyArgGluLe 2223
Db 7626 CCTTACAGAGAGATGCTTTAAATCCCTAGAAAATCAGACATCACTTCAAAATCTTGAAGA 7685
QY 2223 uCysGluLysMetAspIleMetGluLysGluIleSerValLeuArGLeu----- 2239
Db 7686 AAATGGCAAAAGTCCATATATTAATTTGGAAACAGAGTGTGCTACAACTTGAGAGCATGT 7745
QY 2240 -----MetGlnAsnG 2243
Db 7746 TAGTGCAAGAGACTTAGAATTCACCAAGTGTATTAACAAATMAAAGACATGCCAAGAA 7805
QY 2243 uProGlnGlnGluIleAspAspValAlaGluArGMetAspIleLeuGlnIleuSerArGAsnG 2263
Db 7806 AGGCCAGCTTTGAA-----ACAGAAATGCTTCAAAAGAACATGT 7844
QY 2263 nGluIleGlnGluLeuMet-----GluLysIleSerAlaValTySerGlnGlnHisThrLe 2282
Db 7845 AAACCTACAGAAATATGTTGAAGAAAAGTGTGCT----- 7881
QY 2282 uLeuSerSerLeuSerSerGluLeuGlnLysGluThr-----GluAlaHisLysHisCy 2300
Db 7882 -----GCTCTTGTACGTCAAATCCAACTTGAAGCAGTTGAGAAATATGCAAAATTCG 7934
QY 2300 sMetLeuAsn----- 2303
Db 7935 TCAGATTAATTCAAACATTTCTACAGAACTTGAAGAACAAATATTCAGAAATTAATCA 7994
QY 2304 -IleLysGlu-----SerLeuSerSerThrLeuSerArGSerPheGlySerLe 2319
Db 7995 ACAGAAAGAGAGATGAGTGGGTCAGATATATACAGCATTAACCTTGAGAAATTCAGAAAT 8054
QY 2319 uGlnThrGluHisValLysLeuAsnThrGlnIleuGlnThrLeuLeuAsnLysPheLysVa 2339

Db 8055 ACAAGCCAGAGTGTGTGAATATCATCATAGTTTG-----ATTATGAAAAAGAACAGT 8108
QY 2339 lValTyArgThrAlaAlaValLysGluAspHisSerLeuIleLysAspTyArgLysAs 2359
Db 8109 AGAAATTCGACGAAAAAAATGTTTGTAGAAAAGAAAAGAGCTGTAGAACTACAGAGAGCT 8168
QY 2359 pLeuAlaAlaGluGlnLysArGHisAspGluLeuArGLeuGlnIleuGlnCysLeuGluG 2379
Db 8169 ATTTGAGGCGCATGAGAAAACAGAGACAGAAAGAA----- 8205
QY 2379 nHisGlyArgLysTrpSerAspSerAlaSerGluLeuLys----- 2393
Db 8206 -----AGAAAGAGAGCCCTCAGATGTGTGAAGTCTCAAGACACACTAGACTTAT 8258
QY 2394 -----PheCysGluIleGluPheLeu----- 2400
Db 8259 TCATAGCAATGAGAAAGATGATTTTATATGAACTCGAGGCTTTAGAGCTGAATCACT 8318
QY 2401 -----AsnGluLe 2403
Db 8319 GGCTACCAAGCAGACTTCCAGTTATAAAGAAAAGCTGAAAACCTTCAGAAAGAGCT 8378
QY 2403 uLeuPheLysLysAlaAsnIleIleGlnSerValGlnAspAspPheSerGluValGlnVa 2423
Db 8379 TTTGGTAAAGAAACAAAT--ATGACATCTCTTCAGAAAGACTTAAAGCAGTTAGGA 8435
QY 2423 lPheLeuAsnGlnValGlySerThrLeuGln-----GluG 2435
Db 8436 TCACCTCGCAGAGAGCAAAAGAGAAATTCGCATTTTAAAGAAAAGAGATGAGCTGAGGT 8495
QY 2435 uLeuGlnHisLysLysGly----- 2441
Db 8496 ACAAGAAAGCAAAAAGCCGTGATGTTGAGCCACTTCTATTAACCTGATAGAGCAT 8555
QY 2441 ----- 2441
Db 8556 TGATCCACAGACAGATGGAGCTGAAAGATCAGTACACCATCAGACTCCAAATTTCT 8615
QY 2441 ----- 2441
Db 8616 TGTTAAATATCAGAGATACAAATTAATTTACAGAGTGAATGTCTCCAGAAAGACTTAC 8675
QY 2442 -----PheMetGlnThrLeuGlnGluIlePheGlyAspLeuHisValAspAl 2456
Db 8676 TGAATATATCAGTCACTTACTGAAAAATTTGAGAAATGCAAGAACTACATCT-- 8730
QY 2456 aLysLysLeuSerGluGlyMetGlnGlnGluAsnArGArgIleAlaSerThr----- 2473
Db 8731 -----GCTGAATTTTGGACATGGAATCCAGACATATTTCCAGAAACTGAAACCTT 8780
QY 2474 -----IleGlnLeuLeuThrLysArg-----LeuLysAl 2483
Db 8781 AAAGAGGAGAACATATGTGCGCTTCAGTTACTGAAAGAGAAATGTGTGATCTTGAAGGC 8840
QY 2483 aValValGln-----SerLysIle----- 2489
Db 8841 AGTATATCAGTGTGTGAGAAAGTAAAGAGGATCTCAATTCGAGACTACACATTTCTGA 8900
QY 2490 -----GlnArGLeuIleThr-----Va 2495
Db 8901 TGCTTACAGACTAGAGAAATATGCTCCAGTATTCGTATCAGACTGGGGTCAGGAGAT 8960
QY 2495 lTyLeuAsnGln-----PheGluAlaLysLysGlnGluLysGlnGlnAsn-- 2511
Db 8961 TTATCTTACACAGCAGTCAAGGATTTGACATPAGCATCAAGAAAGCCGAGAGAAAGAACT 9020
QY 2512 -----LysGluLeuMetArGArgMetGluHisH 2521
Db 9021 AAGTCACAAGATTCCTTTCCAAAGAAATTAAGAGATTACTGAGACTGTGCATATATGA 9080
QY 2521 sGlyProSerAla-----SerVa 2527

```

Db 9081 AGGATGACAGGTGCTTCTCTCAGCTCCCTATAGTAGTGAAGACCACTTCTAT 9140
Qy 2527 lmetglu-----glugluasnalaargyleuuglylleuleuysrh 2541
Db 9141 TCAGCAGGTTTCAGACCTTGCGCTAGAGAGAAAGCTTACATCAATCACTCATC 9200
Qy 2541 tvalglinspgluserlyslsleuuglnserargilleuysmetleuugluasnleuvas 2561
Db 9201 TCTAAGGATTTAATTACAAAGATGCAA-----CTGCAAAAGAAAGCCGA 9245
Qy 2561 nleuvallysaspsalalamehis-----lysglylulysvalal 2575
Db 9246 GGTATTATGATGTTCTCAATCTCAAGAGAGCTTCAGACTGCGAGGTGAACTACTGCT 9305
Qy 2575 alleuuglnaspysleuau----- 2582
Db 9306 TGCCCTTCACAGTTTCTTTCAGAGAGCCCTAGTCTTTTACTAGACAGATTTCCGACGA 9365
Qy 2583 -----SerArgasnalaqlualagluleuasnalametglinvalysle 2597
Db 9366 GGTGACAGCTCTAGTACTACAGATGAGAGTGGTTTACTAAACTTTGGAAACAGAAAT 9425
Qy 2597 uThrlslysglnaspsalnleuglnalalamehis----- 2609
Db 9426 ACAAGAACAGGGTGTGAATATCAAGCAGCTATGGAATGCCCTCCAGAAAGCAGATAGAAG 9485
Qy 2610 -----gluileguasnleuuglnlysmetvalalalsygslylavalproty 2625
Db 9486 GAGTTTGTATTCTGAATTTCAAGGCACTGCATGCACAAATGAATGAGAAATTAAGTCT 9545
Qy 2625 rlysgluuglualeap-----Asnleuysrh 2634
Db 9546 GAAAGAGAACAGAGAGAGTGAAGAACCAAGCAAGACTCTTGCAATTAATATACACA 9605
Qy 2634 rlyse-----valvalyslleugluuuglystlelytyrserlysal 2649
Db 9606 GAACAGCTCTCAAAATGCTGAGATGCAAGTGAAGTCAAGTGAAGAAA--GACAGAGC 9662
Qy 2649 athraspdlngluilealatyrlleuysercysleuuglnaspyslysgluuglyleuar 2669
Db 9663 AACGGAA-----CTGCAGACAGCTGAGTCTGAGAAATGGTGTGTC 9707
Qy 2669 garleuuglysgluuueuarargalaglinalaspsasnspthrhtrvalcysvalpr 2689
Db 9708 TGAAGTGAAGAGTGAAGTGCACAAACTAAATGGAACTAGAACACACATCAAGCACA 9767
Qy 2689 olyasprytrglnlyslaserthrpherovalthrcysglyglyserglylleua 2709
Db 9768 GCATTAACACCTAATA----- 9783
Qy 2709 lglinserthralmetleuvalleuuglnserglulysalalaleuugluaruglyleuse 2729
Db 9784 -----GAATTCGAGGCTTTACAGCTTGGAAATTAA 9812
Qy 2729 thistyllyslslystyrhishlsleuuserargthmetserseerigluasparg-- 2748
Db 9813 AGATAAGACAGATGACATGATTCCTTAATGACACATTAACAGTGAACGAATAATC 9872
Qy 2749 -----Lyslysthlysalalysseraspsalalhissersehi 2761
Db 9873 AAGAGAGCTCCAGTGGCTTTGGAGAAAGAGAAAGCCCAAG----- 9912
Qy 2761 sthngllysersehihsargglyserpronthlysthrghlythyrarhishgly--Pr 2780
Db 9913 -TTGGGAGCGAGTGAACACGGGATTAAGAAAGAACTTGAAGATTCGAAGTTTCACTGA 9971
Qy 2780 ovalthrprogluargserglumetproserleuhsleuuglyserprolyslserser 2800
Db 9972 GAGTCAGAAACAAAGAAATCTTCAGCTAAATCTACTTTGGAAACACAGAAACAACTACT 10031
Qy 2800 uSerSerthrlslyarvalyalseerproasnargserglulietyrserrglnleuvalme 2820
Db 10032 GAACGAATCCCAAGAAATAAGATACACAGAGATGCTATATGATGCCAGGTGTCTAGA 10091

```

```

Qy 2820 tserprogllysthrnglymethishlslyshlsleuuserproserlysalglyleuhi 2840
Db 10092 AGAACAGGTCGAAACTTAAGACCTTCAAGTACTTCTTCATGATGGAAGATTCGAATTCG 10151
Qy 2840 slslysarvalaleuserproasnargserglumetprothrglnhlsvalilleserpr 2860
Db 10152 GGAATGAGTACTACCTCATGATAGGAGCGGGA----- 10185
Qy 2860 oglysthrnglyleuhslysalasnleuthrghlyser-thrleupheaspsalnleuser 2880
Db 10186 -----TTGCAGCAGCAGCTGCAGACAGATGATGATGACTGCAGACTTCGCGCC 10232
Qy 2880 eProCyslysglnin 2885
Db 10233 ACCCTTGCCCTCAGAGG 10249

RESULT 11
US-10-171-311-1
: Sequence 1, Application US/10171311
: Publication No. US20030087270A1
: GENERAL INFORMATION:
: APPLICANT: Schlegel, Robert
: APPLICANT: Chen, Yan
: APPLICANT: Zhao, Xumel
: APPLICANT: Monahan, John
: APPLICANT: Kamatkar, Shubhangi
: APPLICANT: Glatic, Karen
: APPLICANT: Gannavarapu, Manjula
: APPLICANT: Hoersb, Sebastian
: TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
: TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
: FILE REFERENCE: MRI-035
: CURRENT APPLICATION NUMBER: US/10/171,311
: CURRENT FILING DATE: 2002-06-12
: PRIOR APPLICATION NUMBER: US 60/298,159
: PRIOR FILING DATE: 2001-06-13
: PRIOR APPLICATION NUMBER: US 60/298,155
: PRIOR FILING DATE: 2001-06-13
: PRIOR APPLICATION NUMBER: US 60/335,936
: PRIOR FILING DATE: 2001-11-14
: NUMBER OF SEQ ID NOS: 238
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1
: LENGTH: 12462
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-171-311-1

Alignment Scores:
Pred. No.: 3,19e-56 Length: 12462
Score: 1022.50 Matches: 707
Percent Similarity: 36.54% Conservative: 655
Best Local Similarity: 18.97% Mismatches: 1128
Query Match: 6,92% Indels: 1237
DB: 9 Gaps: 153

US-09-150-867-1 (1-2954) x US-10-171-311-1 (1-12462)
Qy 70 Argseralaleuuglnlytyrasnnglythrilephealatyrglyglnthrsersegly 89
Db 51 CGATCGGCGAGGAG-----CGCGGAGACATGCTTCACTTGGCGGGGAGCCCGGA 104
Qy 90 lysthrtyrthmetmetcglythrproasnserleugly-----lleileproglin 106
Db 105 -----CCGAATCGGCTCTCTAGCGCTGAGCTTGGCGGTC 140
Qy 107 Alaileglnlualpheylystleileglnlulieproasnargglupheuleuarg 126
Db 141 CACCTCGCT-----CCAAATCGACCTTCTCTTCTATC 173
Qy 127 Valserlyr-----Metglulietyr-Asnoluthrvallysasprleuleucysasps 144

```

```

Db 174 CCCAACCCCTCACCCTGTTTCCCTGCCCTTTCGACAGAGCCATGGAGAGCA 233
QY 144 PARGARGLYLSPROLEUGLULIARGGLU-----AspPheasnArgAsnVa 160
Db 234 GGAGAGACAGAAAGAGTGGAGCGCGCAAGAACCTTGCCAGCTTTCGACAAAGAAA 293
QY 160 ITRYVALALASPLEUTHRELUGLULIEMETVALPROGLUHISVALILEGINTRPII 180
Db 294 AGCTCACTCGGATGGCAGAGTCCT-----TC 320
QY 180 ELYSLYGLYLULYSASNARGHISTYRGLYLUTHRILYMETASNAPHSISERSEAR 200
Db 321 CAAAGACAGAAAAAAGAGA-----AAACCTCAAGCAGTAACATGATGTGTC 371
QY 200 GSERHISTHRLIHPHARGMETILEVALGLUSERARGSPARGASNPROTHIRANSE 220
Db 372 AGCACACCAGATTTGATTCATCAATCAGACTGAATGAATGTACATTAATAGTTC 431
QY 220 RGLUASNOCYSAPGLYALVALMETVALSERHISLEUASNLEUVALSPLEUALAGLYSE 240
Db 432 TCAGAGAGTAGATCAACTGTGATTCCTGATCTACATATATGAACTACTACTAGTGG 491
QY 240 RGLUARGALASERGLNTHRLIYALAGLUGLYVALARGLEULYS-----254
Db 492 AGAATATACCATCATGAGCAGGCTTCCTCTGGAACCTGGAAGTGAATTCACACAC 551
QY 255 -----GLUGLYCYSASN-----ILEASNARGSERLEUPHEILEUGLULI 271
Db 552 AGCAGATGACAGGAGTTCAGAGTGAATGGTTGCAATTTCTGTGATG-----597
QY 271 SLYLEUSERASPGLYNALAGLYGLYPHEILEASNTRYARGSPSERLYSEUTHRAR 291
Db 598 -----598
QY 291 GILEUGLINSNSERLEUGLYGLYASNALALYSTHVALILEECYSTHRLIETHRPR 311
Db 618 TTTATTTAAGGAGAGAAATTTGGT-----642
QY 311 OVALSERPHEASPGIUTHRLEUSERTHLEUGLNPHALASERHRLALALYSVALAR 331
Db 643 -----GTTATGATKTTTATTCT-----GACACAGAGACACA 674
QY 331 GASNTHRPROHISVALAENGLUVALLEUASPAPGLUALALEUULYSARGTRYARGLY 351
Db 675 AGACAGTCCGACTCATCTAGAGATGATGAAAGTGGCTGGAGAGCAGCAT-----729
QY 351 SGLUULELEUASPLEULYSGLYNLEUGLINSLEUGLINSERSESERGLUTHRILYSAI 371
Db 730 -GAGATTGAAAGCTAAACAGAGAGCTGGAAAGATGAGGTTACTATGGGACTGAAGG 788
QY 371 A-----GlnAlaMetAlaLysGluLuhIs-----ThrGl 381
Db 789 ACTGACAGAGTTACAAAGATTTGAAGTGCATTAAACAAGAGATGSCATTATTAACCA 848
QY 381 NLEULEUALAGLULIULYSGLINLEUHISLYSGLUARGLUNSPARGILETRPHISLEUTH 401
Db 849 GCTCACTCTATTTACAAACAGAGAAAGATGAGAAAGATGAGAAAGATTTT 908
QY 401 RASNLIEVALVALASERSEGLINUSERGLINLINSAPGLNARGVALYLSARGLYSAR 421
Db 909 AGAGTTG-----ACAGAACAGAGTCAAAAATTTACAGATTCAATTT 947
QY 421 GARGVALIHRTRPALAPROGLYLYSILEGLINSNSERLEUHSIALASERGLYVALSERHS 441
Db 948 TCAGCAATATACAGCTAGTGAANACTGAGAAACAGACACTATATACACAGCTGCAGA 1007
QY 441 P-----PheASPmetLeuSerAr 447
Db 1008 CTTTACTACAGCCAAACACAGATCTCATCTCATCAACAGCAGCTTGAGAACAGACCA 1067
QY 447 GLEUPROGLYASNPHESERLYS-----454

```

```

Db 1068 CTTATTAGAAGATTATCAGAAAAAGAAAGACCTTCACAAATGATTTCTTGCA 1127
QY 455 -----LYSAlaLysPheSerASPmetProSerPheProGluILEASPserValLysTh 473
Db 1128 AGAGAAATTTAAAGTATATGAAATGGACACAGATTAAAAAGTAGAAACCTCAATTAAGA 1187
QY 473 RGLUPHESERASPHEASPALALEUSERMETELASPSEARGNYLILASPALAGI 493
Db 1188 AGAAATATACAGAAAGAGAGCAATCATGAAATTAACACAAAAATATAGAGAAAGA 1247
QY 493 U-----TRPASNLEUALASERLYSVALPHRHSARGGLYLSTh-----SerLe 508
Db 1248 AAAGAAACCTTTGACTTAAGGATTAATTAACACTGCTGATTAATTTCTAGGAGAAATT 1307
QY 508 UHISGLINSEMETILEASPHEGLYGLIN-----ILESERAS 520
Db 1308 ACAAGAACAGATGTGTCCAAAGACCAAGAAATTAATAACATGAAATTAGAGCTGACTAA 1367
QY 520 PSEVALGINPHENHISASPSESERLYSGLUASNGLINLEUGLINTYRLEUPROLYSAPSE 540
Db 1368 TTCTAAGCAAAAAGAAAGACAGCTTCTGAAAGAAATTAACACAGTTAATG-----1416
QY 540 RGLYASPMEVALAGLUCYSARGLYS-----ALASERPHEGLULYSGLULI 555
Db 1417 -GGACAGCTGGAAGAACTTCAGAAAGAAATCATTAAGACAGCAGCTTGGAAACGTGATAT 1475
QY 555 ETHRSEr-----LEUGLNLINGLINSERLYS-----GLUGLULYLYSGLUENVA 573
Db 1476 AGTACAACGAAATGCAACAAACAAACAAAGAAAGTTAGAACCACTCGGGCAGCTGCA 1535
QY 573 LGINSERPHEGLULIULYSTHLEALAGLUGLINSERLYS-----SERVALYSAI 591
Db 1536 TGAGATGTATGGCAGCAGTACTGCAATGAAACAGAAATTAATTAAGACACACATGGC 1595
QY 591 ALYSASNLEUGLUMETVALTHRASNSEARG-----GLUHISSErILEASNALAGI 608
Db 1596 ACAGATGAGAGAAATGAAACACAGGCAATAGGAGAAATGAGAAATGCTTAAGCTCATA 1655
QY 608 UVALGINTHRASPVALGLULYSGLUVALVALARGLYSGLUMESERVALLEULYSPSE 628
Db 1656 TTCAATATTTACAGTTATATAGATCAGATTAAG-----TTAATTTAGATG 1700
QY 628 RGLYTYRASNALASERASNSErASPleUGLINSAPSERSEVALASPGLYLSARGLEUSE 648
Db 1701 GGCATTAATTAAGTAAATTAATTAATTCAGAACTACTCAACAAAGAAACCTCA 1760
QY 648 RSErSErHISASPGLUCYSILEGLUHISARGLYSMETLEUGLINSILYSLIEVALSPLE 668
Db 1761 GGAAGACTAGGACTAATTTTAGAAGAAAGTGTCTCAGAGAGACAGCTTGAAGACCT 1820
QY 668 UGLUGLUPHEILEGLINSLEUASNLYSLYSErGLUASNAPSLYSGLINSERSErGI 688
Db 1821 TGTGGAAGAAATGAGCTTTTCAAGGAGACAGATTTCAGAGAGCTAGACAGCAATAGCTGA 1880
QY 688 UGLINASPHEMET-----GLUSErILEGLINUCYSGLUAlaILEMETAl 703
Db 1881 ACAAGAAAGTAACTTAATGAAACACATATAGTCCCTTAGTACAGTGGAAATTTGAAGC 1940
QY 703 AGLU-----704
Db 1941 TGAGATTGTTTCTGATCTGATTCAGAAAGAAAGAACTGAATTAACATGACAGAGAGT 2000
QY 705 -----LYSAlaASNAlaLeu-----709
Db 2001 TACAATTTACAAGATTAACCTTGAATGTTAGAAAAAGAAAGAAATGCTGTGTAGACAG 2060
QY 709 -----709
Db 2061 AATGCTGAATCACAAGACCTGAATTAGAGAGCTGAGAACACAGCTTCATTATTACTCA 2120
QY 710 -----GLUGLULIEMALALEUWELARGASPASNPHESAP-----ASNILEI 723
Db 2121 CGAAGAAAGAGCTTCCAAACGAGAGAGATTTAGAAATTTGACATGCAATTAATTTGA 2180

```

Oy 723 eleuglunsgluThrlu-----LysArgluIleAlaaspLeuArgse 739
 Db 2181 AAACTTAAGATATATTAGCACTTCACTATTAACAGAGATAGCGTTTACAGAAAGA 2240
 Oy 739 rleuylsglunsgluThrluAsnGluPhegluIleuylsgluThrluGlnlygl 759
 Db 2241 AATGATCAAAAAGATAGAACCATGCACTT-----GAAAGGACAAATTGATAAC 2291
 Oy 759 uHlsgluAlaGlnleuIleHlsgluIleGlySerleuLysLysLeuValGluAsnAlaGl 779
 Db 2292 TAAACAGATCAATTAATTTTGAAATTTCAAGCTAAAGATTACGCGAGTCT----- 2346
 Oy 779 uMetLysAsnGlnAsnLeuGlu-----AspLeuGluThrluLysTh 793
 Db 2347 -CTTGTAATTCAGACAGACAGAAATGACTCTTCAATCAATGAACTTCAAAAGAAAT 2405
 Oy 793 rLysleuLeuylsglu-----GlnGluIle 801
 Db 2406 TGAATTAAGTCAAGACAGAAAGAAAGGATGACTTGACAGAAAGATTCAAGAAAT 2465
 Oy 801 egluLeu-----AlaGluLeuHrGluArgAlaAspAsnLeuGlnly 815
 Db 2466 ACAACTTAACAAAGAAATTTGAGAAACAGATGAGAAAGAAAGAAATGATCTTCAAGA 2525
 Oy 815 sLysValArgAsnPheAspLeuSerValSer---MetGluAspSerGluLysLeuGlygl 834
 Db 2526 AAAATTGCGACAACTTGACGACAGAAATAGCACTTCTTAAGATGCAAAAGAAACCCCTTGA 2585
 Oy 834 uGluIlePheGlnleuLysGlnSerLeuSerAspAlaGluAlaVal----- 849
 Db 2586 AGACATGTGAAATTAATCTCTGCTGTAGCCAGAAAGAAATGATGATTTCTTAGACTC 2645
 Oy 850 -----ThrArgAspAla-----GlnLysGluLysSerPheLeuArgSerGln 863
 Db 2646 CATTAACTCAAAATCCAAAGACTCTGTGTGGAAAAAGAAATAGAAATACCTTAATAGAGA 2705
 Oy 863 uAsn-----LeuGluLeuLysGluLysMetGluAspThrSerAs 876
 Db 2706 AAATGAGAGACTCAAAACAAATGATATTCACCTTAATCAAGATGTAAGAAACCAAGAA 2765
 Oy 876 nTPTrLysAsnGlnLysGluLysAlaAlaSerleuPheGlnLysGlnleuGluThrluGly 896
 Db 2766 CACTTTTCACTTCTCTGAAATAAACTTTGAAGTAACTATCAAGAGATTACAAAGAGAGTA 2825
 Oy 896 sSerAsnTyrLysLysMetGluAlaAspLeuGln-----LysGluLeuGln 911
 Db 2826 TGCCTTGCCTTCTCAAAAGTAAAGATGATTTAGAAAGACAGTAAATAATAACAGAAATAGA 2885
 Oy 911 n-----SerAlaPheAsnGluIleAsnTyrLeuAsn----- 921
 Db 2886 GTATTAAGAACTTAAGACACTTAATGAGAGCTTCATTTCGCAAGAAATTAATCCAAC 2945
 Oy 921 ----- 921
 Db 2946 TACAGTGAATAGAAAGTCTGTCTTGATGAGAGACAAACTTTTGTAGCAGAAACAT 3005
 Oy 922 -----GlyLeuLeuAlaGlyLysValProArgAspLeuLeu----- 933
 Db 3006 GGAATGGGTGAGGTGTTGAAAGAGATACAAAGACACTCATGAAAACTTGAGGTAAAC 3065
 Oy 934 -----SerArgValGluLeuGluLysLysValSerGluPheSerLysGlnleuGlnly 951
 Db 3066 CAAGCAGAGAAATTAAGCTGCTACAGAGACTGTCTGATCTTTCTGCAACAAATTTGAAACA 3125
 Oy 951 sAlaLeuGluGlnLysAsnAlaLeuGluAsnGluValThrCysLeuSerGluLysLys-- 970
 Db 3126 GAAACATGCTGAGATGATTTCTTAATGAGAGATTAATCTTTAAAGCAAGAAAGAA 3185
 Oy 971 -----PheLeuProAsnGluValGluCysLeuLysAsn----- 981
 Db 3186 ACAAGTTTCATTGAGATGTAGAGAGCTAGAAATCATTAATTAACCAACAAAGGAGCAAGAAA 3245

Oy 982 -----GlnIleSerLysAlaSerGluGlnIleMetLeuLeuGlygl 995
 Db 3246 TGTACAGTCTGTGATTAAGTCAAGTAAAGCTTTTATTAGATGAGATGTTGACCATCAAG 3305
 Oy 995 nGluGly---GlnHlSerAlaSerLleSerLys----- 1006
 Db 3306 CAGGGGTCTGAAAGGATCAGTTCTTAAGTAAATTAAGTTTGGTGAAGATCAAAAT 3365
 Oy 1007 -----GlnGluIleIleMetGlnGlnSerGluGln-- 1017
 Db 3366 AATGGTGAAAGTAAAGTTCTTTGAAATATGAGCTGTGGAGAAAGTAAAGTAAAGTAA 3425
 Oy 1018 -----IleLeuGlnleuThrAspGluValThrHlArgInsar-----LysValGln 1033
 Db 3426 ACAGTTGATTTTGATCACTTACATCTGTAAACAAAGAAATCATCTTAAGACAACTCA 3485
 Oy 1033 nGlnThrGluGlnGlnTyrLeuGlu-----MetLysLysMetHlAspAs 1048
 Db 3486 ACCAAGTGAAATGATTAACCTCAGAAAGAACTCAATGATTAATTAATGAGAACAGATGA 3545
 Oy 1048 PheuPheGluLysTyrIleArgAsnLysSerGluAlaGlu----- 1061
 Db 3546 TTTA-----AGCTACAGATGAGAACCCCAACCAATTCGCTCTCTGTGT 3590
 Oy 1062 -----AspLeuLeuArgGlu---MetGluAsnLeuLysGlyThr----- 1073
 Db 3591 TTATTCACATCATGTGATCAGGTTCGTGAATATATGGAATTAAGAAATTAAGCTCT 3650
 Oy 1074 -----MetGluSerValGluValLysIleAlaAspThr 1084
 Db 3651 TTGCAGCTTAAGAAAGACTTATTTTGTCTCAAGAGAAAGATCAAGAACTTCAGAA 3710
 Oy 1084 rLysHlsgluLeuGlu---GluThrIleArg-----AspLysGluGlnle 1098
 Db 3711 AATACACAGATTGAATACAGACTATGAAACACAAAGAAACAGGTGATGAAGAAAGCC 3770
 Oy 1098 uLeuHlS-----GluLysLysTyrPheP 1106
 Db 3771 TTTACATCTGCTCATGGAAGAACTCAAAAGCACTGTGAGAAATGTCTTATTTT 3830
 Oy 1106 egluAlaMetGlnThrIlePheProIleThrProLeuSerAspSerLeuProPterLys 1126
 Db 3831 ACAGACTTATGACAGTGC-----CTTGCTGAATATTTATACCTCGCTT 3875
 Oy 1126 sLeuValGluGluAsnSerGlnAspProIleGluIleAsnAspLys----- 1141
 Db 3876 AAATGTGAAGTAAATGAGAGACAAAGACAAATTCGTGTATTAATCTGAAATTA 3935
 Oy 1142 -----HisAs 1143
 Db 3936 AGATCCAAATTAACAAGATTATAGATAGAACTTCAAGACTTTCAAAGAAATATATGCACAC 3995
 Oy 1143 nLeuIleAlaLeuAlaThrGluArgAsnAsnIleMetValCysLeuGluThr----- 1160
 Db 3996 TCTTCTCAACAAAGTAAAGAAATACAAACAACTCTGTGACTTCAACAGCAGATAG 4055
 Oy 1161 -----GluArgAsnSerLe 1165
 Db 4056 CAAGATCTGGGACAGACAGACAGATGATGAACCTGAATTTGGAAGAAACCTTCC 4115
 Oy 1165 uLysGluGlnVal-----IleAs 1171
 Db 4116 AAAGAGCAACAGAGTTTATTAATCAATTCATCTCAATGACCAATTTGGAAGACATTTGA 4175
 Oy 1171 PheuAsnThrGlnleuGlnSerLeuGlnAlaGlnSerIleGluLysSerAspLeuGlnly 1191
 Db 4176 TGTCAATCATTAAGCAAGATATCTTCTGCAAGATCTTGAAGAAACCTTAACCTGAGAA 4235
 Oy 1191 sProLysGlnAspLeuGlu-----GluGlygl 1200
 Db 4236 ACAAGTTCAAGAAATTAAGAAAGCTCATATCTCTTGTGACAGCAACAGTTGAAGAAAGTGA 4295
 Oy 1200 uValLysLeuLeuLeuGluMetGluLeuLeuLysGlyHlSLeu----- 1214

QY	1777	gaenleuasnnglunwctulmwetvalmwetleuglunwctuglunleuylsasnsergl	1797
Db	6306	AGAGTGGGAAGAACACACTGACTAGG---TTTAAAGAGCGTGGAAACAAAATAATATGCA	6362
QY	1797	nargthValillealagluaragspclinleuglnaaspaerleuarxgluserValglume	1817
Db	6363	ACTAATG-----	6369
QY	1817	tserllegluThrlhAspaerleuarxglysalaglnalaleuglnglnglnlysas	1837
Db	6370	-----GATTTAAGACAGCAAAACCAACGACATTGGAAAGCAGTTGAGA	6410
QY	1837	plysValaglnleuThrserlnlleser-----Valle	1849
Db	6411	AAAAATAGAAAATTTTATGATGACGACGCATTGACAGAGAACATGAGAGATGTATT	6470
QY	1849	uglnglnlysalleserleuenglunsnnglnmetleuTyraenValalathrVallysgl	1869
Db	6471	CCAACTGAGAAATACAGAACTAGAACAGCACTT-----AAGGTGTTCCTCGATTCCA	6524
QY	1869	uThrleuSerclunArgaspaerleuasnlnserlysglnHlsleuPleserGlillegl	1889
Db	6525	GCCATATAGTAA-----CATCAAACTAGAGAGTTGGA	6557
QY	1889	uThrleuSerleuSerleuysgluysglu-----Phealale	1902
Db	6558	ACAGTTGCAAAATCCTGTAAATAAATAAACAACAATGCAATGACATGAGCTTTGTCTCTCA	6617
QY	1902	uglnglnalaglnAspIysAspIysAlaAspAlaAlaArgIysThrlleAspIleThrgIly	1922
Db	6618	AGACACACTTCCAAGGGATTTATACAAAGAAAGAAAGAAATAGAAAGAAAGAGAGTTGAC	6677
QY	1922	silesAsnillegluglnleuenglnclnalathraenleuylsgluThrleuTy	1942
Db	6678	AGTAAGAGAACTGAGCAGCGCTTCTTGATGATGCAGATCTTTTCAAAAGCTAGAGA	6737
QY	1942	rgluArgglu---SerleuileglnCysIysgluGlunleuAlaleuasnThr-----Gl	1959
Db	6738	CCGAAAACACTTGGAGCTGTATACAACTTAACACGAAATGTCTCTAGAGATCAATTGCA	6797
QY	1959	uHlsleuArggluThrleuylsSerlyspAspIysleuAlaleuGllysmetGlnglnglu-	1978
Db	6798	GCGTGAACGAGATGCCATATACACAGAAAGGAAAGAGATTACAACTTGAAGAGCAATT	6857
QY	1979	-----ArgAspIunAlaAlaasnlysalillealaleuThrgIlysmetSerse	1995
Db	6858	AGAACACTTATAGAGAACTGAAATATAGATGAAGAAGTTCAACAATTACATATGCA	6917
QY	1995	rleuglnglnglnlleasnlnuasnValhThrlreuylsgluGluglngIygluylsgl	2015
Db	6918	ATTAGAAATACAAAAAAGGAATCTACGCCGCTTACAGAACTTGAACAGGAAACAA	6977
QY	2015	uThrPheTyrlleuglnArgProserleuysglInserSerseIymetGlnglngluAr	2035
Db	6978	ATTATT-----AAGATGACATGAGGAACACTGG	7007
QY	2035	ggluSerleuylsThrlYAspIleuglnglnleuglnglnlalegu-----	2049
Db	7008	ACTTGGCATTAAGAAATCTGATGCCATGTCTACTACAGCCACCAATGCTATTGGGA	7067
QY	2050	-----LysgluIleSerGluIathrasngIuIleuylsAsnleuThrlalysIlese	2067
Db	7068	ATTTGGCTCAAAATATACAGGAAAAAGAGTAAATATGCAATTAAATGAACAATTAC	7127
QY	2067	rSerleuGluglnglnleuGlunAlasnAlaserlleuasnlnuAlaIsergluAr	2087
Db	7128	GAATCTGCACAGCAACTT-----AAATTTCAACAGATATCAAGGTATTGAAGAA-	7179
QY	2087	ggluAsnleuArgHlsSerlysglnglnleuValsergluenglunleuSerleuTh	2107
Db	7180	-----AAAAATCTGTATTAAGGAGATCTTTTGAACCCAAATAGATG	7220

QY	2107	rlEulysSerArghsrhlsalArPhealagInSerLyArVgGluLyAsrGluAlaValAs	2127
Db	7221	TTTTGGATGGGTGATCAGAAATGTGTGAAGAAATAAGAAAGAAATAGACAGCTCA	7280
QY	2127	nlvslleAla-----	2130
Db	7281	TGAAAGTGAATTAAGAAACTTCAACAGCAATTGGCAAAATATTGGACAGAAACATCAATGA	7340
QY	2131	-----SerleuAlaGluGluIleLysIleLeuThrLysGluMetAspGluPheArgAs	2148
Db	7341	TGCTCATTCCTCTCAGAGAAAGACAGACATTTAAAAACATCAATTGAGATGCTGTTATAGC	7400
QY	2148	pSerLysGluSerLeuGluGlnGlnSerSerHlslEusSerGluGluLeuCySthTrLyL	2168
Db	7401	TGAAAGAGCTGGCTTGGAAACAGACAGCAAGTAGAAACCGCTAATGAAAGAAATGCCTTCATGAA	7460
QY	2168	sThr-----	2174
Db	7461	AAATGTACTTAAGAAACCAATTTTAAATGAATTCAGTACACAGGAATTTTCAGCTT	7500
QY	2174	ulYsGlnGlnLysGluLysAspIleAsnAsnLysleuAlaGluLysValLysGluValAspG1	2194
Db	7521	AAAGAGAAACAGCTGAAGTGTG-----GAAAGATTCAAAGATACACAGA	7555
QY	2194	u-----LeuLenglnHlslEusSerLeuLysGluGlnLeuAspGlnI1	2209
Db	7566	GAATAGTGTATACGTGGCTGTAGATCATCTGCAGCAAGCAAAACCTGACATCTGAGT	7625
QY	2209	eglMetGlu-----LeuArGAsnGluLys-----LeuArGAsnTrpGluLe	2233
Db	7626	CTTACACAGAGATGCTCTTAATCCCTAGAAATTCAGACATCTTCAAATCTTTTGAAAG	7685
QY	2233	uCySglLysMetAspIleMetGluLysGluIleSerValLeuArgLeu-----	2239
Db	7686	AAATGCAGAAAGCTTCATATATATTGGAAACAGAGTTGCTACAACTGAGACAGACGT	7745
QY	2240	-----MetGlnAsnG1	2243
Db	7746	TAGTCAGAAAGCACTTAGACCTTACCAGTGTATTAACAATAAAGACATCAAGAAACA	7805
QY	2243	uProGlnGlnLulLysAspArAlaAlaGluArMetAspIleLysGlnSerArGAsnG1	2263
Db	7806	AGGCCAGTTTAA-----ACAGAAATGCTTCAAAACAGACATTTGT	7844
QY	2263	nglulLleGlnLysLeuMet--GluLysIleSerAlaValTrpSerGluGlnHlslThrLe	2282
Db	7845	AAACCTACAGAAATTACTTGAGAAAAAGCGCTGT-----	7881
QY	2282	uLeuSerSerLeuSerSerGluLeuGlnLysGluThr-----GluAlaHlslLysHlslCy	2300
Db	7882	-----GCTCTGTGCAGTCAAAATCCAACTTGAGCGAGTTCAGGAATATGCCAAATTCTG	7934
QY	2300	sMetLeuAsn-----	2303
Db	7935	TCAGATATATCAAAACAATTCATCAGAACTGAAGACAAATATTACAGATTTAAATCA	7994
QY	2304	-lLleLysGlu-----SerLeuSerSerThrLeuSerArGserPheGlySerLe	2319
Db	7995	ACTAAGAGAAATATAGTGTGGGTGCAGATATATCAGCATTAACCTTGAGAAATTCAGAAAT	8034
QY	2319	uGlnThrGlnHlslValLysLeuAsnThrGlnLeuGlnThrLeuLeuAsnLysPheLysVa	2339
Db	8055	AGAAAGCCAGGTGTGAAATTCATAGTACTAGTTG-----ATTTTAGAAAAGAACAGT	8108
QY	2339	lValTrpArgThrAlaAlaValLysGluAspHlslSerLeuIleLysAspTrpGluLysAs	2359
Db	8109	AGAAATTCAGAAAAAATATTTTAAAGAAAAAGAAAGAGTGTAGAAATACAGAACT	8168
QY	2359	pLeuAlaAlaGluGlnLysArgHlslAspGluLeuArgLeuGlnLeuGlnLysLeuGlnG1	2379
Db	8169	ATTGGAGGGCAATAGAAAAACAGACAGAGAAAGA-----	8205
QY	2379	nhlG1ArGTrpSerArGserAlaSerGlnGluLeuLys-----	2393

Db 8206 -----AGAAAAAGAACCCCTCAAGATGTTGAGTTCTCAAGACAACACTAGCTATT 8258
 QY 2394 -----PheCysglutIleGluPheLeu----- 2400
 Db 8259 TCATAGCAATGAGAAAGTGGATTTTATGAACTCGAGGCTTAGAGCTGAATCACT 8318
 QY 2401 -----AsnGluLeu 2403
 Db 8319 GGTACCAAGACAACTTGCCTATTAAGAAAAGGCTGAAAACTTCAAGAAAGAGCT 8378
 QY 2403 uLeuPheLysLysAlaAsnIleIleGlnSerValGlnAspAspPheSerGluValAlaIle 2423
 Db 8379 TTTGCTAAAGAAACAAAT---ATGACATCTCTCAAGAAAGCTTAGCCAGGTAGGGA 8435
 QY 2423 lPheLeuAsnGluValGlySerThrLeuGln-----GluG1 2435
 Db 8436 TCACCTGCGAGAGCAAAAAGAAATGTCCATTTTAGAAAAAGAGATGAGACTGAGGT 8495
 QY 2435 uLeuGluHisLysLysGly----- 2441
 Db 8496 ACAAGAAAGCAAAAAGGCTTCATGTTTGAGCCACTTCTATAAACTGAGTAAGAGCAT 8555
 QY 2441 ----- 2441
 Db 8556 TGCATCCAGACAGATGGGACTGTGAAGATCAGTAGCAGCAATCAGACTCCACAATTC 8615
 QY 2441 ----- 2441
 Db 8616 TCTTAAATGCAAGATACAAATTAATTACAGATGAATGTTCTCCAGAAAGATTAC 8675
 QY 2442 -----PheMetGlnTrpLeuGluGluPheGluLysPheHisValAspAl 2456
 Db 8676 TGAATTAATCACTCAGTTACTGAAAAAATTTGAGAGAGATGCAAGACACTACGTCT 8730
 QY 2456 aLysLysLeuSerGluGluMetGluGlnLysAsnArgArgIleAlaSerThr----- 2473
 Db 8731 -----GCTGAATTTTGGACATGGAATCCAGACATATTTCAGAAACTGAAACCTT 8780
 QY 2474 -----IleGlnLeuLeuThrLysArg-----LeuLysAl 2483
 Db 8781 AAGAGGAACAACATATGTTGCCGTCAGTTACTGAAAGAGAAATGTGTTACTTCAAGGC 8840
 QY 2483 aValValGln-----SerLysIle----- 2489
 Db 8841 AGTGATACAGTGTCTGAGAAGTAAGAGGATCTCAATTCTCAGTAGACATTCGTA 8900
 QY 2490 -----GlnArgGluIleThr-----Val 2495
 Db 8901 TGCCTACAGACTAGAGAAATATGCTCCAGTAGATTCTGATCAGACTGGGCTCAGGAAAT 8960
 QY 2495 lTyrlLeuAsnGln-----PheGluAlaLysLeuGlnGluLysGluGlnAsn-- 2511
 Db 8961 TTATCTTACACACAGTCAGGAGATTGACATAGCATCAGAAAGCCGAGAGAAAGTA 9020
 QY 2512 -----LysGluLeuMetArgArgMetGluHisI 2521
 Db 9021 AAGTGAACAAGATTCTTCCAAAAGAAATTAAGGATTACTAGAGCTGCTCATATATA 9080
 QY 2521 sGlyProSerAla-----SerVal 2527
 Db 9081 AGGATGCAAGTGTCTTCTCTACTAGTCTCCCTATAGCATGAGAGAGACCATCTCAT 9140
 QY 2527 lMetGlu-----GluGluAsnAlaArgLeuLeuGluGlyIleLeuLysTh 2541
 Db 9141 TCACAGAGTTTCAGAACCTTGGCTAGAGAGAGAAAAAGCTTACATCAATACATCTCAGC 9200
 QY 2541 rValGlnAspGluSerLysLysLeuGlnSerArgIleLysMetLeuGlnLysLysLeuAs 2561
 Db 9201 TCTTAAGAGATTTAATTAACAAGATGCAA-----CTGCAAGAGAGAAAGCCGA 9245
 QY 2561 nLeuValLysAspAspAlaMetHis-----LysGluGluLysValAl 2575

Db 9246 GGTATATGATAGTTCTCATCTCATAGAGCTTCTCAGACTGGCGAGGTGAACACTGCT 9305
 QY 2575 aLleLeuGlnAspLysLeuLeu----- 2582
 Db 9306 TGCCCTTCACAACTTTCTTAGAAGAGCTAGTGTCTTACTAGCAGACATTTCCGAGGA 9365
 QY 2583 -----SerArgAsnAlaGluAlaGluLeuAsnAlaMetGlnValLysIle 2597
 Db 9366 GCTGACAGCTCTAGTACTACAGATGCGAGTGTCTTACTTAAGCTGTTTGGACAGAAAT 9425
 QY 2597 uThrLysLysGlnAspAsnLeuGlnAlaAlaMetLys----- 2609
 Db 9426 ACAAGAACAGGGGTGTGAATATCAAGCGACTATGGAATGCCCTCCAGAAACAGATGAAAG 9485
 QY 2610 -----GluIleGluAsnLeuGlnLysMetValAlaLysGlyAlaValProTy 2625
 Db 9486 GAGTTTGTATCTGAATATTCAGGCACCTCATGACCAATATGATGTGAGAAATTAAGTCT 9545
 QY 2625 rLysGluGluLysP-----AsnLeuLysTh 2634
 Db 9546 GAAAGAGAACAAAGAGAGCTGAGAAACCAAGCAAGACTTGGAAATATATATACAGCA 9605
 QY 2634 rLys-----ValValLysIleGluMetGluLysIleLysTySerLysAl 2649
 Db 9606 GAAGCAGCTCAATATGCTGAGATGCAAGTGCAGCTCAGCAGTATGAAA---GACAGAGC 9662
 QY 2649 aThrAspGlnGluIleAlaTyrlLeuLysSerCysLeuLysAspLysGluGluGluLeuAr 2669
 Db 9663 AACGGAA-----CTCAGAGACAGCTGAGTTCTGAGAAATAGTGCTGCTGC 9707
 QY 2669 garGleuLysGluGluLysLeuArgAlaGlnAlaAspAsnAspThrValLysValPr 2699
 Db 9708 TGAAGTCAAGAGAGTACCTTGCACAAACTTAATTTGAACTAGAAACACCAAGAGCACA 9767
 QY 2689 oLysAspTyrlGluLysAlaSerThrPheProValThrCysGlyGlyLysGlyIleVal 2709
 Db 9768 GCATTAACACCTTAAA----- 9783
 QY 2709 lGlnSerThrAlaMetLeuValLeuGlnSerGluLysAlaAlaLeuGluArgGluLeuSe 2729
 Db 9784 -----GATTCGAGGCTTCAGGTTGGAAGTTAA 9812
 QY 2729 rHisTyrlLysLysTyrlHisLysLeuSerArgThrMetSerSerSerGluAspArg-- 2748
 Db 9813 AGATTAACAGAGATGATCATTTGCTTAATGACATTAAGCAAGATGCAAGAAACAAATC 9872
 QY 2749 -----LysLysThrLysAlaLysSerAspAlaHisSerSerHis 2761
 Db 9873 AAGAGACTCCAGTGGCTTGGAGAAAGAAAGCAAG----- 9912
 QY 2761 sThrGlySerSerHisArgLysSerProHisLysThrGluThrTyrlArgHisGly--Pr 2780
 Db 9913 -TTGGAGACGAGTAAAGACGGGATTAAGAAAGAACTTGAGATCTGAAGCTTTTCACTTGA 9971
 QY 2780 oValThrProGluArgSerGluMetProSerIleHisLysGlySerProLysLysSerG1 2800
 Db 9972 GAGTCAGAACCAAGAGATCTTCAGCTTAATCTACTTTTGGAAACACACAGAAACAACTACT 10031
 QY 2800 uSerSerThrLysArgValAlaSerProAsnArgSerGluIleLysSerGlnLeuValMet 2820
 Db 10032 GAGCAATCCACGCAAAAAAATAGAAATCAACAGAAATCTATAGATCCCAAGTTGCAGA 10091
 QY 2820 tSerProGlyLysThrGlyMetHisLysHisIleLeuSerProSerLysValGlyLeuHis 2840
 Db 10092 AGAACAGGTGCAAACTTAGACTTCAGGTAATCTTGAATCTGAGAAAGTTGCAATTCG 10151
 QY 2840 sLysLysArgAlaLeuSerProAsnArgSerGluMetProThrGlnHisValIleSerPr 2860
 Db 10152 GGAATTAAGTATACCTTACAGTATAGGAGCCGGAA----- 10185
 QY 2860 oGlyLysThrGlyLeuHisLysAsnLeuThrGluSer-ThrLeuPheAspAsnLeuSer 2880
 Db 10186 -----TTGCACGCAAGACTGTCAGAGCAGATGATGTACTGACAGCTCGGCC 10232

Db 5803 TGTGACCTCCAGAGAGGTGACTCTCTCAGCAAGAGCAGAGCTGCTGCAGAAAGC 5862
 QY 680 ----- 680
 Db 5863 CAGGAGAGAGCTCGCGCAGAGCTGGAGTGTCTAGACAGAGAGGATGGCCCTGGCAAG 5922
 QY 681 ----- 681
 Db 5923 GGTAAATGTGAGCTTCAGCTGCAGAGGGGAGACTCTGCCAGGCGCCAGAGAGAGAGCAGCA 5982
 QY 692 TGlusertleuLeucys-----GluAlaIleMet----- 702
 Db 5983 GGAGGAGCTGCAGCTGCTGCTGCGGAGAGGAGAGCTTTCAGAGAGATGCTGATGGCCCT 6042
 QY 703 -----AlagluLysAlaAsnAlaLeuGluLeuAlaLeuMetArgAspAsnPhaAsp 721
 Db 6043 GGAAAGCCAAACAGTCACATACACTCAGTGAACATGATCATCTTCGGGAAAGCCCTGGAGTC 6102
 QY 721 nilelleuLysngluLysnleuLysArgGlu-----IleAl 734
 Db 6103 AAGTCACCTGGAGAGGGGAGTACTAGAGCAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 6162
 QY 734 aspleuLysArgSerleuLysGlu-----AsnGluLysnleuLysnleuLysnleuLys 751
 Db 6163 TAGGGCAGAGAGTCATATTGCAAGAGCTGCTGAGTTCTGAAACACCTGGAAGACAGAGT 6222
 QY 751 eleuLys-----LysGluThrGlnLysGluHisGluAlaGluLeuIleH 766
 Db 6223 AGCTATCTTCGGGCTGCAGCTGCAGCTGCAGCTGCAGCTGCAGCTGCAGCTGCAGCTGCAG 6282
 QY 766 sGluIleGlySerleuLysLysLeuValGluAsnAlaGluMetArgAsnAlaAsnLeuG 786
 Db 6283 TAAATTTGGCTGAACAGCAGACTTCTCCAG-----TTAGAGAGAGAGAAACAGCTGTGTG 6339
 QY 786 uGluAspleuGluThrLysLysLeuLysGluGluLysGluLysGluLysGluLysGluLys 806
 Db 6340 CAGCAGAAATGGAGGCGGAGCAGAGCGAGAAATGCTTTCAGAGCTGCAGCTGCAGAGG 6399
 QY 806 uArgLysAlaGluAspAsnLeuGluLysValArgAsnPhaAspLeuSerValSerMet 826
 Db 6400 AGAGAAAGAGG----- 6409
 QY 826 tGlyAspSerGluLysLeuGluGluLysLeuPheGluLeuLysGluSerLeuSerAspAl 846
 Db 6410 -----AGGAAAGCCCTGTGGGAAAGAAACACTCAGCTGCAGGCTGCAGAGAAAGC 6462
 QY 846 agluAlaValThrArgAspAlaGluLysGluLysSerPheLeuArgSerGluAsnLeuG 866
 Db 6463 TGGAGAGGCTGGGCTGAGCTGCAGCTGCAGCTGCAGCTGCAGCTGCAGCTGCAGCTGCAG 6522
 QY 866 uLeuLysGluLysMetGluAspThrSerAsnTrpTyrAsnGluLysGluLysAlaAla 886
 Db 6523 AATTCAAAAGAACTAAGTACAGTCACT-----CACCAAGAGAGAGAGAGAGAGAGAG 6573
 QY 886 rLeuPheGluLysGluLeuGluLysThrGluLysSerAsnTrpTyrLysMetGluAlaAsp 906
 Db 6574 TCAGCTGAG-----CAGCTTATCATCAGAGAGCAGAAAGCAGAGAGAGAGAGAGAGAG 6630
 QY 906 uGluLysGluLysLeuLysSerAlaPheAsnGluLysAlaGluLysGluLysGluLysGlu 926
 Db 6631 AGTCCAG 6690
 QY 926 yLysValProArgAspLeuSerLeuSerArgValGluLeuGluLysLysValSerGluPhe 946
 Db 6691 C---GTGAGAGGTGAG 6747
 QY 946 rLysGluLeuGluLysAlaLeuGluLys----- 955
 Db 6748 GGAGCTACTGAG 6807
 QY 956 -LysAsnAlaLeuGluLysGluLysValThrCysLeuSerGluTyrLys---PheLeuProAs 974
 Db 6808 CAAAGGGCAGCTGGAGAGTCCAGATTCAACAGTGCATCAAGCCAAAGAGAGAGAGAGAG 6867

QY 974 nGluValGluLysLeuLysAsnGluLysSerLysAlaSerGluGluLysMetLeuLeu 994
 Db 6868 GGAAGTGAAGTCCCTGAGAGCTGGAAGTGCAGACACGAA-----CG 6906
 QY 994 sGluGluLysGluHisSerAlaSerIleIleSerLysGluGluLysMetGluGlu 1014
 Db 6907 GAGTCAAGCAG 6960
 QY 1014 nSerGluGluLysLeuGluLysLeuThrAspGluValThrHisThrGlnSerLysValGlu 1034
 Db 6961 GCMAAG 7020
 QY 1034 nThrGluGluGluLysLeuGluLysMetLysLysMetHisAspLeuPheGluLysTyr 1054
 Db 7021 GCTCCGGAGAAATGGAG 7060
 QY 1054 eArgAsnLysSerGluAlaGluAspLeuLeuArgGluMetGluLysnleuLysGlyThrMet 1074
 Db 7061 -----CAGGAGCTGCAGAAAGCTCTGCAGAGCTTGAAGAAAGGAGAA 7101
 QY 1074 tGluSerValGluValLysIleAlaAspThrLysHisGluLeuGluGluLysThrIleArgAs 1094
 Db 7102 AATGAGCTGGAATGAGAGCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7161
 QY 1094 pLysGluGluLysLeuHisGluLysLysTyrPhePheGluAlaMetGluThrIlePhePr 1114
 Db 7162 GAGGAGAAAGAAAG 7210
 QY 1114 oIleThrProLeuSerAspSerLeuProSerLysLeuValGluGluLysnSerGlnAs 1134
 Db 7210 ----- 7210
 QY 1134 pProIleGluLysAsnAspTyrHisAsnLeuIleAlaLeuAlaThrGluArgAsnAsn 1154
 Db 7210 ----- 7210
 QY 1154 eMetValGluLeuGluLysThrGluArgAsnSerLeuLysGluGluValIleAspLeuAsnTh 1174
 Db 7211 -----CTGGAAACAGAG-----MAGGAGAGAGTCCCTCTGCAGAGAC 7248
 QY 1174 rGluLysGluSerLeuGluAlaGluSerIleGluLysSerAspLeuGluLysProLysG 1194
 Db 7249 ACTGCTCAG 7308
 QY 1194 nAspLeuGluGluLysGluValLysLeuLeuGluMetGlu-----LeuLeuLys 1211
 Db 7309 GGACATG-----AAAGTCCAGAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7359
 QY 1211 sGluHisLeuThrAspSerGluLeuSerIleGluLysLeuGluLysGluLysGluLysGlu 1231
 Db 7360 GAGCCAGCTCAG 7411
 QY 1231 lThrGluLysLeuGluLysThrLeuGluGluLysGluLysGluLysGluLysGluLysGlu 1251
 Db 7412 -AGAGATGACCTTCTCCCTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7470
 QY 1251 uLeuGluLysnleuPheGluAspLeuLysAlaGluHisAspSerLeuLysGluLysLeu 1271
 Db 7471 CCTGCAG 7524
 QY 1271 rGluAsnIleGluLysSerIleGluThrGlnAspGluLeuArgAlaAlaGluGluLys 1291
 Db 7525 CCAGAGAGCTGTGGAGAG-----GAGGTTCCAGAGAACTGAGAGAGAGAGAGAGAGAGAG 7581
 QY 1291 uArgGluLysGluLysGluLysValAspSerPheArgGluGluLysLeuLysPysSerVal 1311
 Db 7582 CCGAATTCAGAGAGAGAGAG-----GAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7623
 QY 1311 yLieserProAsnHisAspAlaValAlaAsnGluLysValSerLeuGluGluVal 1331
 Db 7624 TCTCTCAGTGAAGAGAGAGAG-----CAGAGAGCTCCTGTGTTTAAAGAGAGC 7671

QY	1331	lanserleuglnserclu-----metleuargclgcluatgaspcluleuglnth	1348
Db	7672	TGACCTATTTCGACCAACAAGAGCTGAGTCCCTCGCCAGGCATGACGAGGCCAGG	7731
QY	1348	rsercylgslalevalsercluleuglnleueuathgalahtslvalyservalgl	1368
Db	7732	AGAACAAAGAGACTGAGTCTAGTGGATTTACTTAAAGCAAGAGTGAAGAAAGGA	7791
QY	1368	uglgluasleugluilethrlyslsleuansnglueuglulysleuglyly	1388
Db	7792	GGCTGACTTTCGTG-----GCCAGGAAGACAGTGTCTGGAGGA	7830
QY	1388	sserglugluserglvalleulyssermetleuglunsleulysgluasprasnly	1408
Db	7831	GCTGGAGGCTTCATATCCAGGACAGCAGCTG-----CG	7866
QY	1408	sleulysgluglaleaglucluluryserserlysglsuansglnpheserleugluva	1428
Db	7867	AGCCTCTTTGGGCCCCAGGAAGCCAGGACGCCCACTACAGCTGCCACTGGCCAGAC	7926
QY	1428	lpheserlyserglnlylsleuvalaspluilegualleulysalaglnleulysal	1448
Db	7927	AGAGAGC-----CAGCTAGAAGCGTGGCCGACAGACAGACGC	7965
QY	1448	alalagluclul---Argleugluileulysaspargsplyrphlegluvalglnthral	1467
Db	7966	CGGGAACAGCCAGCCAGCCAGCCAGCCAGCTGCACGCTTACTGTCCCTGCCAGAGC	8025
QY	1467	asanthrasleuvalgluclylysleuglnthrproleuglnalasphts-----	1484
Db	8026	CTTGGGTCTGTTTGAGAGAGCAGGCCCTAGCTGAGTGGGGAGACTGTCTCTTC	8085
QY	1485	-----gluglnasperlleaspargaryserserleuglunmetgluleulysva	1500
Db	8086	CGTCTGGGCGCTTGAAGCAGACCAAGATGAGACTGAGAGC-----	8125
QY	1500	lleuglyglulysleugluarvasnlnlrytleuglnuargleuglnlglulysle	1520
Db	8126	-----CTCTTTTAAAGAGAGGCCCCCTGTGACTGTCTCTCGCTGAGGACGT	8172
QY	1520	ugluuser---Asnlylsleu---gluileuglnlsglunetglunthrservalle	1538
Db	8173	AGCATCTGCCCTCCACAGCTTATGAAACCTGTGGAGACTCAACAGACCCGGATGT	8222
QY	1538	uleulysaspaspaleuglnlinsleuglnserleueusergluasnlleilelely	1558
Db	8233	TCTGAGGAGTCAAGTCC---CAGAACTGGAAAGAGGCTCACTAGTACTGAGGCTGA	8289
QY	1558	sgluasnlleasprthrthleulyshtslsersasprthrglnalaglnleuglnlysth	1578
Db	8290	GAGCAGAGCTCCACAGAGTTTCAG-----GATCGAGAGACAGCTCTCCCGAA	8340
QY	1578	rglnglngluleuglnleualalysnsleualllelaalaaseraasprnlyproil	1598
Db	8341	TCAGGAAGAAATCTAAG-----	8359
QY	1598	ethnglnglulysglulthseralaspys---Valhlsproleuglnlglulyslele	1617
Db	8360	-TGGGAAGAAACACAGACTCCCTTAGAACCTGATGAGAACTACATGAAACTATAGGC	8418
QY	1617	uleulysleucluleuhtslslnlysthansngluneglunlulysleuhtslslyly	1637
Db	8419	ATCTTCACAGAGTCCGCTGGGAGACACAGACTGACGAATGGAAGCCAGGGGTGAGCG	8478
QY	1637	sansngluleuglnlaleaglnvalgluleulyscysgluvalglunhsleuhtslsle	1657
Db	8479	AGAGTTACTCTCAGAGCCCAAGAGAACTGTGACAGCCAGGTGGAACTCGCAAGACGC	8538
QY	1657	rmetlleugluserlysserserleucluserleuglnhtslslnlysthsprrthnglu	1677
Db	8539	TGTCTGTAAGAACCCAGGGCTCAGGCAAGTCTCTGTGGCATCTGGAAAGAACTCGAGAC	8598
QY	1677	nglnleuvalaleulysglunlmetglunlvalalthrnglnlulysglu-----	1695

Db	8599	GGCTGGCTCAGACACTGAGGTGAAATAATGGAGAGTAGAGAGTGAACGGTGAAGACGCCA	8658
QY	1696	----LeuGlnGlnThrHisGluHisLeuThrThrIaGluValaAspHisLeuLysGluAsnI1	1714
Db	8659	GGCTCTGCACAGACAGGCGCACTGAAAGTGGCCCAAGGGAAGGCGCTCTGCACAGAAATT7	8718
QY	1714	eglLeuGluLysLeuAsnPhelysAsnGluIaGlnGlnLysThrLysGluGlnCysLe17	1734
Db	8719	GGCCCTC-----CT	8727
QY	1734	uLeuAsnGluAsnLysGluLeuLeuGlnInserserGlnHisAlyLeuGlnCysGluIleGlu1	1754
Db	8728	GACCCAGACCCCTAGCTTAAGAAAGAGAGAGGTGAGACTCTCGGGGCAAAATCCAGGA	8787
QY	1754	uLeuMetLysSerLeuLysAspLysGlnSerAlaIaLeuGlnThrLeuLysGluSerGlu1	1774
Db	8788	ACTGAGACAGACAGACGGGAATGCGAAGAGCTGTTTGAAATTCGTCTCTGAGACTGAA	8847
QY	1774	nLysValIleAsnLeuAsnGlnGluMetGluMetValMetLeuGluMetGluLeuLeu1	1794
Db	8848	GAAG-----AGCAACCCAGAGGTAGATCGACGAAGAAAGATTCACAGACTAGA	8898
QY	1794	sAsnSerGlnaArgThrAl-----IleIaGlnIaArgAspGln-----	1806
Db	8899	GAAGTAGAGCTCTGTTTAAAGCATCTGCCATGGCCGATCGCCAGACCGAGACCAAGACT	8958
QY	1807	----LeuGlnAspAspLeuArgGlnSerValIaGlnMetSerIleGluThrGlnAsp--	1822
Db	8959	GACTCTGCACAGAGGAGACAGATCAGAGAG--CTCGAGAAAGATCGGAGACTCAGAGGA	9015
QY	1824	-----AspLeuArgLysAlaGlnGluIaLeuGlnGlnGlnLys1836	
Db	9016	CGTCTTGAGACATACGCTTCTTAGACCTTGAGAAAGAACCAAAATGATTGATCCCAAG	9075
QY	1836	sAspLysValaGlnGluLeuThrSerGlnIleSerValLeuGln---GluLysIleSerLe1855	
Db	9076	AGCAGACGTTCTAGGACTCGAAAAAGCAGATGGTTACTCTGGAATGCCCTGGCAACT	9135
QY	1855	uLeuGlnAsn-----GlnMetLeuLysThrAsnValaThrValThrLysGluThrLeuSerG1	1873
Db	9136	GGAGGAAAAACCATCACAAGATGGAGTGCACGCAAAAACATGATCTCAGAGACTGAGGCCA	9195
QY	1873	uArgAspAspLeuAsnGlnSerLysGlnHisLeuPheSerGlnIleGluThrLeuSerLe1893	
Db	9196	GAGGGAACCCAGAGAGTGGCTTTGACCCACTTACGCTGAGACCTAGAAAGAGACCA	9255
QY	1893	uSerLeuLysGluLysGluPheIaLeuGlnGlnIaGluLysAspLysAlaAspAlaI1913	
Db	9256	GGACCTCTCAGGCACAAAGACGACGATCCATGCACTGGAGAGCCACACAGCACCGTTCTGGC	9315
QY	1913	aArgLysThrIleAspIleThrGlnLysLysIleSerAsnIleGlnGlnGlnLeuLeuGlnG1	1933
Db	9316	AAGGAGACTCGAGAGAGGAGCACAGGAGGGAAGTCTCAGCGAGAACAGATC-----	9367
QY	1933	naIaThrAsnLeuLysGluThrLeuLysArgLysArgLysSerLeuLysCysLysGluG1	1953
Db	9368	-----GAGGA	9372
QY	1953	nLeuAlaLeuAsnThrGlnHisLeuArgGluThrLeuLysSerLysAspLeuAlaLeuG1	1973
Db	9373	GCTGCAAGGACGACAAAGAGCATCTGACTCAGATCTCGAGAGAGAGACAGAGACTG--	9430
QY	1973	ylYsMetGlnGlnGlnArg-----AspGluAlaIleAsnLysValIleAl1988	
Db	9431	ATGCTCCAAAGAGAGAGGATTCAGGTCTTCGAGATTCAGAGACCCGGCACCAACAGAT	9489
QY	1988	aleuThrGluLysMetSerSerLeuGlnGlnGlnIleAsnGlu-----AsnValThr	2005
Db	9490	CTGTGAGAGAGACTGGAACAGATCAAGCTGTCTTGAAGAGAGCGAGCCGGGAGCTAAC	9549
QY	2005	rThr-----LeuLysGlnGluGlnGluLysGluLysGluThrPh	2017

Db 9550 CACTCAGAGCAGCTGATCAGGAAAGGCGCAGAGGAAAGGCGC----- 9595
 QY 2017 eTyrLeuGlnArgProSerLysGlnInSerSerSerGlnMetGlnLeuArgGluSe 2037
 Db 9596 -----CCAGTAAAGCAGCAGCGCGGAGC-----CTAGACACATGAAAGCTGAT 9639
 QY 2037 rLeuLysThrLysAspLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 2057
 Db 9640 CCTGGCTGTAAG 9699
 QY 2057 nGlnLysAsnLeuThrAlaLysIleSerSerLeuGlnGlnGlnGlnGlnGlnGln 2077
 Db 9700 GCTCAAGACCAAGCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 9759
 QY 2077 aSerIleLeuAsnGlnAlaValSerGlu-----ArgGlnAsnLeuArgHisSe 2093
 Db 9760 CCT 9819
 QY 2093 rLysGlnGlnLeuValSerGlnLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 2113
 Db 9820 CAGGGAACAA-----GGGAGCTGAAAGAGCAGTCACTT----- 9853
 QY 2113 sAlaIlePheLeuGlnSerLysArgGlnLysAspGlnAlaValAsnLysIleAlaSerLeuAl 2133
 Db 9854 -----CAGAGTCAACTGTGATGAGCCCAAGAGCCCTAGCCCAAGAGGA 9897
 QY 2133 aGluGlnLysLysIleLeuThrLysGlnMetAspGlnPheArgAspSerLysGlnSerLe 2153
 Db 9898 CCAGAGACCTGGAGGCTCTGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 9957
 QY 2153 uGlnGlnInSerSerHisLeuSerGlnGlnGln-----CysThrTyrLysThr 2169
 Db 9958 GAAGAAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 10017
 QY 2169 r-----GluLeuGlnMetLeuLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 2186
 Db 10018 GCGTCATGAGAGCTTCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 10077
 QY 2186 aGluLysValLysGlnValAspGlnLeuLeuGlnHisLeuSerSerLysGlnGlnLe 2206
 Db 10078 AGTGGAGGAGCGCGGCT-----CAGGCCCTGAGGAGGAGGCTCT 10116
 QY 2206 uAspGlnIleGlnMetGlnLeuArgAsnGlnLysLeuArgAsnTyrGlnLeuGlnGln 2226
 Db 10117 GAGAGACTAAGGGCTGAGTCTCG----- 10141
 QY 2226 sMetAspIleMetGlnLysGlnLysSerValLeuArgLeuMetGlnAsn-----GluPr 2244
 Db 10142 -----GAACAGGAGAAAGCTCTGTGGCCCTCAGCAGCAGCTGTCTGAGCA 10188
 QY 2244 oGlnGlnGlu-----GluAspAspValAlaGlnArgMetAs 2256
 Db 10189 GGCACAGAGCATGATGAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 10248
 QY 2256 rLeuLeuGlnSerArgAsnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 2276
 Db 10249 AGTCTCAAGAGAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 10294
 QY 2276 rSerGlnGlnHisThrLeuLeuSerSerLeuSerSerGlnGlnGlnGlnGlnGlnGln 2296
 Db 10295 -----AGTCAAGTCTCTCCCGCATCAGAGAGGAGGCTGCGCCG 10329
 QY 2296 aHisLysHisCysMetLeuAsnIleLysGlnSerSerSerThrLeuSerArgSerPh 2316
 Db 10330 GCGCCCG-----GCTGAGGCTCTCAGAGAGCCCTTGGCAGAGCTCA 10371
 QY 2316 eGlySerLeuGlnThrGlnHisValLysLeuAsnThrGlnLeuGlnThrLeuLeuAsn 2336
 Db 10372 TGTGCGCTCAGAGGAGAAAG-----CAGCATCTCTC----- 10405
 QY 2336 sPheLysValValLysArgThrAlaAlaValLysGlnAspHisSerLeuLysAspTyr 2356
 Db 10406 -----GAGCAGGAGAAATTTAGCGCGCACTCT 10431

QY 2356 rGlnLysAspLeuAlaAlaGlnGlnLysArgHisAspGlnLeuArgLeuGlnGlnGln 2376
 Db 10432 GAGGCCAGCACTGCAACC-----CTGCAAAGCTCTCCCTGATGC 10470
 QY 2376 sLeuGlnGlnHisGlnArgLysTrpSerAspSerAlaSerGlnGlnLysPheCysGln 2396
 Db 10471 CTGCCAGGCACACAGCTGGCAG-----CTGAGAGAGGCTCTGAGAGATCAAGA 10518
 QY 2396 uIleGlnPheLeuAsnGlnLysLeuPheLysLysAlaAsnIleIleGlnSerValGlnAs 2416
 Db 10519 AGGTGAGATCCAGAGCCAGCATCTC-----CGATACCAAGA 10554
 QY 2416 pAspPheSerGlnValGlnValPheLeuAsnGlnValGlySerThrLeuGlnGlnGln 2436
 Db 10555 GATGTGACAGCAGCTGCGACGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 10602
 QY 2436 uGlnHisLysLysGlnPheMetGlnTrpLeuGlnGlnPheGlnLysPheuHisValAspAl 2456
 Db 10603 GAGACATCAGCAGAGAGGAGCAGCAGCTGTGAGAG-----TCTCTGCC 10647
 QY 2456 aLysLysLeuSerGlnGlnMetGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 2476
 Db 10648 CCAGAGGCTCCAAAGAGATATGATCAAGAGACAGATCTGGCCAGAGAGAGAA-- 10705
 QY 2476 uLeuThrLysArgLeuLysAlaValGlnSerLysIleGlnArgGlnLysThrValTyr 2496
 Db 10706 -----GAGAGCAGATTAAGGGCTTCACTCAGAGT-----GTAAAGAGCTACAGCTGAC 10755
 QY 2496 rLeuAsnGlnPheGlnAlaLysLeuGlnGlnLysLysGlnGlnAsnLysGlnLeuMetAr 2516
 Db 10756 TCTAGCCCAAAAGAGAGAGATCTGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 10815
 QY 2516 pArgMetGlnHisGlnGlnProSerAlaSerValMetGlnGlnGlnGlnGlnGlnGln 2536
 Db 10816 GGAAGCCTTACCCACAGCCACAAACCTCCCAATGAGAGAGAGAGAGAGAGAGAGAGAG 10875
 QY 2536 uGlyIleLeuLysThrValGlnAspGlnSerLysLysLeuGlnSerArgIleLysMetLe 2556
 Db 10876 TTTCTTAGAGCCCGAGCTGAGCGGAGCTGAGCGGCTACAGGACCCCTGAGAGAGAG 10935
 QY 2556 uGlnAsnGln-----LeuAsnLeuValLys 2564
 Db 10936 AGAAGCCAGAGAGATGAGTGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 10995
 QY 2564 sAspAspAlaMetHisLysGly-----GluLysValAlaIleLeuGlnAspLysLeuLe 2582
 Db 10996 GACCAAGGCCAGAGTGTGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 11055
 QY 2582 uSerArgAsnAlaGlnAlaGlnLeuAsnAlaMetGlnValLysLeuThrLysLysGlnAs 2602
 Db 11056 GGAAGCGGAGCTCAGAACAGCA-----AGCCTGACAGGA 11088
 QY 2602 pAsnLeuGlnAlaAlaMetLysGlnIleGlnAsnLeuGlnLysMetValAlaLysGlnVal 2622
 Db 11089 TGAAGTGAAGCTCAGCAGAGGAGGCTGTGAGAG-----GAGCGCTACACAGCCAGAGTGC 11145
 QY 2622 aValProTyrLysGln-----GluLeuAs 2630
 Db 11146 AACCCAGACAGCAGAACTGGGGTCCAGAGGAGCAGAGGAGTGTCCAGCTGGAGAGAGTTC 11205
 QY 2630 pAsnLeuLysThrLysValValLysIleGlnMetGlnLysIleLysTyrSerLysAlaThr 2650
 Db 11206 AGGAGTGAAGCTGAGCTAGTCTGATGAAATGAGAGAGAGAGAGAGAGAGAGAGAG 11254
 QY 2650 rAspGlnGlnIleAlaTyrLeuLysSerCysLeuAspLysGlnGlnGlnGlnGlnGln 2670
 Db 11255 -----AGACAAAGGCTTGAACCTGACAGCAGAGAGAGAGAGAGAGAGAGAGAG 11292
 QY 2670 gLeuLysGlnGlnLeuArgArgAlaGlnAlaAspAsnAspThrThrValCysValProLys 2690
 Db 11293 GCTGAGAGATTGACAG 11329

QY 2690 sasPTyrglnLysAlaSerThrPheProValThrcysgllygllySerGlyIleValgl 2710
 DB 11330 -----GTCCAGCTGCGGAGTACCTTGGAG-----CAGGATGGGAGGAGCAGAA 11376
 QY 2710 nserThrAlaMetIleVal-----LeuGlnSerGluLysAlaIleLeuGlnArgGlu 2728
 DB 11377 CTCAGATGCGCAAGTGTGTGCTGCAACTGCAGAAAGAGGTCTCTCCACACTGAGCT 11436
 QY 2728 user---HisTyrlsLysLysTyrlsHisLeuSerArgThrMetSerSerSerGlu 2747
 DB 11437 GACTTTGGAGCGGAGCAGAAAGCAGACGACTACATCACCCGCTGCACAGACCCGCTGA 11496
 QY 2747 pArgLysLysThrLysAlaLysSerAspAlaHisSerSerHisThrGlySerSerHis 2767
 DB 11497 G-----CTAGCAGGCTGCACCA 11514
 QY 2767 gGlySerProHisLysThrGluThrTyrlsArgHisGlyProValThrProGlnArgSer 2787
 DB 11515 CAGCTCTCACACTCTTCTTGGCCGTGGCCAGGCCCT-----GA 11556
 QY 2787 umetProSerLeuHisLeuGlySerProLysSerGluSerSerThrLysArgVal 2807
 DB 11557 GGGCACTGCTG-----GAGGCAGAACCCCGCAGGCTG-- 11590
 QY 2807 IserProAsnArgSerGluIleTyrlsSerGlnLeuValMetSerProGlyLysThrGly 2827
 DB 11591 -----GATGAGTCCCTGACTCAAGTCTGACATCCCGAGGCGAGTCTGCT 11637
 QY 2827 thIsLysHisIleLeuSerProSer 2835
 DB 11638 ACACCCC-----AGCCCCAGC 11653

RESULT 13 US-10-098-841-40

; Sequence 40, Application US/10098841
 ; Publication No. US20020197679A1

GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Liu, Chenghua
 ; APPLICANT: Asundi, Vinod
 ; APPLICANT: Xu, Chongjun
 ; APPLICANT: Zhou, Ping
 ; APPLICANT: Ma, Yundong
 ; APPLICANT: Wang, Jian-Rui
 ; APPLICANT: Zhao, Qiong A.
 ; APPLICANT: Ren, Relyan
 ; APPLICANT: Chen, Rui-hong
 ; APPLICANT: Wang, Duntui
 ; APPLICANT: Wang, Zhiwei
 ; APPLICANT: Wehrman, Tom
 ; APPLICANT: Zhang, Jie
 ; APPLICANT: Qian, Xiaohong B.
 ; APPLICANT: Dmanac, Radoje T.
 ; TITLE OF INVENTION: No. US20020197679A1el Nucleic Acids and
 ; FILE REFERENCE: Polypeptides
 ; CURRENT APPLICATION NUMBER: US/10/098,841
 ; PRIOR FILING DATE: 2002-03-13
 ; PRIOR APPLICATION NUMBER: 09/598,042
 ; PRIOR FILING DATE: 2000-06-20
 ; PRIOR APPLICATION NUMBER: 09/552,317
 ; PRIOR FILING DATE: 2000-04-25
 ; PRIOR APPLICATION NUMBER: 09/488,725
 ; PRIOR FILING DATE: 2000-01-21
 ; NUMBER OF SEQ ID NOS: 331
 ; SOFTWARE: PC_FGenes Version 1.0
 ; SEQ ID NO 40
 ; LENGTH: 6386
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (150)..(5492)

US-10-098-841-40
 Alignment Scores:
 Pred. No.: 1,048-48 Length: 6386
 Score: 900.50 Matches: 457
 Percent Similarity: 38.248 Conservative: 366
 Best Local Similarity: 21.248 Mismatches: 694
 Query Match: 6.108 Indels: 635
 DB: 9 Gaps: 77

US-09-150-867-1 (1-2954) x US-10-098-841-40 (1-6386)

QY 5 AspAlaValLysValCysValArgValArgProLeuIleGlnArgGlnGlnGly----- 22
 DB 318 GATTATCTCCAGGTTGTGTTCGAATAAGCATTATACACAGTCAGAAAAAGAACTTGAG 377
 QY 23 -----AspGlnAlaAsnLeuGlnIleThrLysAlaGlyAsn 34
 DB 378 TCTGAGGGCTGTGTCATATTCGATTCACAGACTGTGTCTGAAAGAGCTCAATGC 437
 QY 35 ThrIleSerGlnValAspGlyThrLysSer-----PheAsnPheAsp 48
 DB 438 ATCTTGCTGCGTTAAAGTCAAAAAGCTCAGGCACATGCGACAGAAATTCAGTTTTC 497
 QY 49 ArgValPheAsnSerHisGluSerThrSerGlnIleTyrlsGlnIleAlaValProIle 68
 DB 498 AAGGTTTGGCCCGACCACTACAGAAAGAAATTCCTTCAGGCTTCATATTCAGAACCA 557
 QY 69 IleArgSerAlaLeuGlnGlyTrpAsnGlyThrIlePheAlaTrpGlyIleThrSerSer 88
 DB 558 GTAAAGACCTCTTGAAGAGCAGAGCTGCTGATTTTACTTACCGGGCTAACCAATTCA 617
 QY 89 GlyLysThrTyrlsThrMetGlyThrProAsnSerLeuGlyIleIleProGlnAlaIle 108
 DB 618 GGAATAACATATCATTCATTCAGGACAGAAAGAAATTTGGCATTCGCTCGAAGCTTGG 677
 QY 109 GlnGluValPheLysIleIleGln-----GluIleProAsnArg 116
 DB 678 AATGTTATTTATGATGATGCTTCAGAAAGAGCTATACAAAGATGAACTTAAAGCAAT 737
 QY 117 -----GluIleProAsnArg 121
 DB 738 AGATCCAGAGAAATCTTAAAGTTATCATCAGAACAGAAAGAAAGAAATTCCTACAA 797
 QY 122 GluPheLeuLeuArg----- 126
 DB 798 AGTGCAATGCTTGCAAAATTAAGAGGTTACTGTGCATTAATGATGATGATCTCT 857
 QY 126 ----- 126
 DB 858 TATGAAAGTTTAACTAAGCTTGAATATCTCAGACTTGAAGATCCATTAAGATTAAT 917
 QY 127 -----ValSerTyrlsMet 130
 DB 918 GAACAAGCCAACTGTAATATGCTTAATTAATTAATTTCTGTGTGCTTCTCTCTT 977
 QY 131 GluIleTyrlsAsnGlnThrValLysAspLeu-----CysAspAspArg 146
 DB 978 GAATTTACAAATGAATATTTATGACTTATTTTCTTCCTGATCATCTAAATCCAAAG 1037
 QY 147 LysLysProLeuGlnIleArgGluAspPheAsnArgAsnValTyrlsAlaAspLeuThr 166
 DB 1038 AGAAAGATGCTGCGCTTCCCAAGCGTAAGGCGTATCTTTTAAAGATCTCAAA 1097
 QY 167 GluGluLeuValMetValProGlnHisValIleGlnIleThrLysGlyGlyLysAsn 186
 DB 1098 TCGATTCAAGTATCTGATTCCAAAGAGCTATAGACTTTTAAACTAGAAATTAAGCAC 1157
 QY 187 ArgHisTyrlsGlyLysThrLysMetAsnAspHisSerSerArgSerHisThrIlePheArg 206
 DB 1158 CAGAGTGTGCTTCACAAATTAATTAATGCTTCAGTAGAAGTACAGCATATTCACCT 1217
 QY 207 Met---IleValGluSerArgAspArgAsnAspProThrAsnSerGluAsnGly 225

```

Db 1218 GTTAAATATTACAGATTGAAGAT-----TCTGAATGTCCTCGT 1256
OY      ::::: |||
226 AlAlaMetValSerHisLeuAsnLeuValAspLeuAlaGlySerGluAlaSerGln 245
OY      ::::: |||
1257 GTAATTCAGTCAGCAGATTTATCTTTATGTCATCTGCTGTTGAGACAACTATGAC 1316
OY      246 ThrGlyAlaGluGlyValArgLeuLeuGlyGlyCysAsnIleAsnArgSerLeuPheIle 265
Db 1317 ACACAGAAATGAGGTGAAAGGTTAAGAGAGACTGGAAATATCAACACTGTTTATGACT 1376
OY      266 LeuGlyGlnValIleLeuLysLeuSerAspGlyAlaGlyIlePhe-----Ile 282
Db 1377 CTGGAAAGTGTATTAACGCTTGTGAAGATAGTGAAGATGCAAAAGTCAACACCATG 1436
OY      283 AsnTyrArgAspSerLeuThrArgIleLeuGlnAsnSerLeuGlyAsnAlaLys 302
Db 1437 CCTTCCCGGAAAGTAACTGACTACTATTTTCAAACTTTTAAATGTAAGGAA 1496
OY      303 ThrValIleIleCysThrIleThrPro-----ValSerPheAspGluThrLeuSerThr 320
Db 1497 ATTTGATATGATTTGCAATATTCAGCAATGTTATTTAGCCTATGATGAACACTCAATGTA 1556
OY      321 LeuGlnPheAlaSerThrAlaLysHisValArgAsnThrProHisValAsnGluValLeu 340
Db 1557 TTGAAGTCTCCGCACTGACAAAAAAGTTGTGTCACAGACACTTTAAATTCCTCTGAA 1616
OY      341 Asp-----AspGluAlaLeuLeuLysArgTyrArg 350
Db 1617 GAGAAATTTATTTGGACCTGCAAAATCTCTCAAGATGTATCTACATGACAGTAATTCAAAC 1676
OY      351 LysGluIleLeuAspLeuLysLysGlnLeuGluAsnLeuGluSerSerGlnThrLys 370
Db 1677 ACTAAATATATTAAATTAAGAAAGCCACCATTTCAATGGAAATATGCTAAGAAATTTG 1736
OY      371 AlaGlnAlaMetAlaLysGlnGluHisThrGlnLeuLeuAlaGluIleLysGlnLeuHis 390
Db 1736 ----- 1736
OY      391 LysGluArgGluAspArgIleThrPheIleThrAsnIleValAlaLysSerGlnGlu 410
Db 1737 ATGGAAACGCGAGATTTGGTTGAGGAGCTAAGAAAC----- 1772
OY      411 SerGlnGlnAspGlnArgValLysArgLysArgArgValThrThrPalaProGlyLysIle 430
Db 1773 GCTGAAGAAACTCAAAATGTGGAAACTAA----- 1802
OY      431 GlnAsnSerLeuHisAlaSerGlyValSerAspPheAspMetLeuSerArgLeuProGly 450
Db 1803 -----CTTCTTGATGAAAGCATCTAGATAAAACATTAGAGAA 1838
OY      451 Asn-----PheSerLysLysAlaLysPheSerAspMetProSerPhePro 465
Db 1839 AATAAGGCTTTCATTAAGCCAGAGAGAGAAAGAAACTGTGTCGACTTA----- 1886
OY      466 GluIleAspAspSerValCysThrGlnPheSerAspPheAspPalaLeuSerMetMet 485
Db 1887 -----ATAGAAAGACTTGAAGAAAAAATCGATAAATGA 1919
OY      486 AspSerAsnGlyIleAspAlaGluThrAsnLeuAlaSerLysValThrHisArgGluLys 505
Db 1920 AAAAAGGAAATTAACCTTGAAATTTAAATTCGAGAGAGATTAACAGAGACTTACT 1979
OY      506 ThrSerLeuHisGlnSerMetIleAspPheGlyGlnIleSerAspSerValGlnPheHis 525
Db 1980 CAGTTTGGCTCAACGGGAAGCTGACTTTAAGAG-----ACTCTGCTTCAAGAGACA 2033
OY      526 AspSerSerLysGluAsnGlnLeuGlnIleThrLeuPalaAspSerGlyAspMetAlaGlu 545
Db 2034 GAGATATTAGAGAAAT-----GCTGAA 2057
OY      546 CysArgLysAlaSerPheGluLysGluIleThrSerLeuGlnGlnLeuGlnSerLys 565

```

```

Db 2038 CCTGCTTGGCTATCTTCAG-----GATTTGTTGGTAAATGTGACACTGCA 2105
OY      566 GluGlu-----GluLysLysGluLeuValGlnSer 575
Db 2106 GAAAGAGCAGCGAAAGACATTTTGCCACAAAGTTGAACTGAAAGAGCTACTGCTGT 2165
OY      576 PheGluLeuLysIleAlaGluLeuGluGluGlnLeuSerValLysPalaLysAsnLeuGlu 595
Db 2166 TTAGAATCAAAATTTAATCAAAATTAACCTGAAATAGCT-----AAACCCAAAGCA---GAA 2219
OY      596 MetValThrAsnSerArgGluHisSerIleAsnAlaGluValGlnThrAspValGluLys 615
Db 2220 TTAATC-----AAACCAAAAGCA 2237
OY      616 GluValValArgLysGluMetSerValLeuGlyAspSerGlyTyrAsnAlaSerAsnSer 635
Db 2238 GAGTTAAAGAGAGAA-----AATGAATCAGATTTCA 2270
OY      636 AspLeuGlnAspSerSerValAspGlyLysArgLeuSerSerSerHisAspGluCysIle 655
Db 2271 TTGATTCAGAGCTTGACATCTAATAAGAAATTAATTAACAGAAATCAAAATTTAA 2330
OY      656 GluHisArgLysMetLeuGluGlnLysIleValAspLeuGluGluPheIleGluAsnLeu 675
Db 2331 AAATGATTAATTAATTAATGTCAAAAGAGATACATCAAGGAATTT---CAGAACCTA 2387
OY      676 AsnLysLysSerGluAsnAspLysGlnLysSerSerGluGlnAspPheMetIleSerIle 695
Db 2388 AAGTCATATGAAAACATTTAAATGCAATTAATGCAAGACAGCTGAT----- 2432
OY      696 GlnLeuCysGluAlaIleMetAlaGluLysPalaAsnAlaLeuGluLeuAlaLeuMet 715
Db 2433 ---ACATCTCTTAAATTAATTAACATTAATTTATTTGTAATGAACAGATTGAAGTACCT 2489
OY      716 ArgAspAsnPheAspAsnIleIleLeuGluAsnGluThrLeu---LysArgGluIleAla 734
Db 2490 AAGGACAGCAAAATTAATTAATCTGTCAGAAAGAAAGAAAGTAAATGAATGAATCTCAG 2549
OY      735 AspLeuGluArgSerLeuLysGluAsnGlnGluThrAsnGluPheGluIleLeuGluLys 754
Db 2550 CAAGATGAAACCCACAGCAAGAAAGAGGCTATATCCATGTTAGTACCTATACATGAAAGAC 2609
OY      755 GluThrGlnLysGluHisGluAlaGlnLeuIleHisGluIleGlySerLeuLysIle 774
Db 2610 CAAAAGAAATGAAAGAAAGTCCGACCGACATTCACAAATTAAGACACAGAGTTTA 2669
OY      775 ValGluAsnAlaGluMetTyrAsnGlnAsnLeuGluLysPheGluThrLysThrLys 794
Db 2670 CAAGAAAT-----AATGAAGAGCTGAGACATTTTATCACTATATGGAAT 2717
OY      795 LeuLeuLysGluGlnGluIleGlnLeuAlaGluLeuArgLysArgAlaAspAsnLeuGln 814
Db 2718 CACTTTAAATTAAGAAAGAGAAAGAAAGCAGATTAATTAACAGATTTGCTCAATTTTCAG 2777
OY      815 LysLysValArgAsnPheAspLeuSerValSerMetGlyAspSerGluLysCysGlu 834
Db 2778 CAG-----GAATCTCTCTTCTGAAAAAAGAAATTAATTAAGTTAA 2822
OY      835 GluIlePheGlnLeuLysGlnSerLeuSerAspAlaGluAlaValThrArgAspAlaGln 854
Db 2823 GAGGTCCAAACAAATTCAGCAAAATTAATGATATGCAATTCGT----- 2864
OY      855 LysGluCysSerPheLeuArgSerGluAsnLeuGluLeuLysGluLysMetGluAspThr 874
Db 2865 ---GAATTCAGTCGAGAAAGTAAATGAAGAACAGGAGAAAGAAATCATGAATG 2921
OY      875 SerAsnTyrPyrAsnGlnLysGluLysAlaAlaSerLeuPheGluLysGlnLeuGluThr 894
Db 2922 TCA-----AATCAGATTAAGAACTGCTACAGACACATTTACAAATATGTTTCACA 2972
OY      895 GluLysSerAsnTyrLysMetGluAlaAspLeuGlnLysGluLeuGlnSerAlaPhe 914
Db 2973 ATAAATTAATGACACAGAAATTAAGACGAACTAGCTGATTCAGTTTCATGAGATT 3032

```

QY .915 AsnGluIleAsnTrpLeuAsnGlyLeuLeuAlaGlyValProArgAspLeuSer 934
 Db 3033 TCAGAACATAGATTGCTCAT----- 3053
 QY 935 ArgValGluLeuGluLysValSerGluPheSerLysGlnLeuGlnLysAlaLeuGlu 954
 Db 3054 -----CTCAGGAGATCTGTCAAATGCTTCAG 3080
 QY 955 GluLysAsnAlaLeuGluAsnGluValThrCysLeuSerGluTrpLysPheLeuProAsn 974
 Db 3081 GAGGATAT-----TTGCCAAT 3098
 QY 975 ---GluValGluCysLeuLysAsnGln-----IleSerLysAlaSerGluLysLeuMet 991
 Db 3099 ACACAGTAGACCTTTAGGATGATATATTGTTGTAAGCAAGATTAAACAAATATGCA 3158
 QY 992 LeuLeuLysGlnGluGluLysHisSerAlaSerIleIleSerLysGlnGluIleIleMet 1011
 Db 3159 ATT-----CAAGAACCCCAATAGGAAATCTTCCACTAGTATGTAAGCTATTGG 3212
 QY 1012 GlnGluGlnSerGluGlnIleLeuGlnIleuThrAspGluValThrHisThrGlnSerLys 1031
 Db 3213 ---GAAAGATGTAAAGAGATTGTGAAGCCTTCCAAAAAGCATCAGATTGAGAA 3269
 QY 1032 Val---GlnGlnThrGlnGluGlnIleTrpLeuGluMetLysLysMetHisAspSer----- 1048
 Db 3270 CTGGAACAACAATTTGAAAAATTCGACGACAGAAAGTAAAGCTATTAAGATGAAACAAT 3329
 QY 1049 ---LeuPheGluLysTrpLysArgAsnLysSerGluAlaGluAspLeuLeuArgLysMet 1067
 Db 3330 AGACATAAGGAGAAAGAGACATTAACCA-----GATGACCTCTATAAAGAAAA 3380
 QY 1068 GluAsnLeuLysGlyThrMetGluSerValGluValLysIleAlaAspThrLysHisGlu 1087
 Db 3381 GAACCTCT-----ATACACACCTGGAAGAAAGA 3410
 QY 1088 LeuGluGlu---ThrIleArgAspLysGluGlnLeuLeuHisGluLysTrpPhePhe 1106
 Db 3411 TTGCAAGAAAAAATGTTACCTTGATGTTCAATACAGCAT----- 3452
 QY 1107 GlnAlaMetGlnThrIlePheProIleThrProLeuSerAspSerLeuProProSerLys 1126
 Db 3452 ----- 3452
 QY 1127 LeuValGluGluLysAsnSerGlnAspProIleGluIleAsnAspTrpHisAsnLeuIleAla 1146
 Db 3453 GTAGTTGAGGAAG-----AGA 3470
 QY 1147 LeuAlaThrGluArgAsnAsnIleMetValCysLeuGluIleuThrGluArgAsnSerLeuLys 1166
 Db 3471 GCGCTTCAGAACTTACACAAAGGTGTACTGC-----TATTAAG 3509
 QY 1167 GluGlnValIleAspLeuAsnThrGlnLeuGlnSerLeuGlnAlaGlnSerIleGluLys 1186
 Db 3510 GCAAAAATAAGGAACTTGAACA-----ATTTCAGACTCACAAGAACTGGAAGCT 3560
 QY 1187 SerAspLeuGlnLysProLysGlnAspLeuGluGluGluGlu-----ValLysLeu 1203
 Db 3561 AGTATTCAGCCAACTTGAACAAGACATTTTGGAAAAGAAATCTATCATCTTAAAGCTA 3620
 QY 1204 LeuLeuGluMetGluLeuLeuLysGlyHisLeuThrAspSer-----Gln 1218
 Db 3621 GAAAGAAATTTGAAGGAATTTCAAGAACATCTTCAGATCTGTGCAAAAACACCAAAAGAT 3680
 QY 1219 LeuSerIleGluLysLeuGlnLeuGlnLysLeuGluValThrGluLysLeuGlnIleu 1238
 Db 3681 TTAATATGAAAGCACTCACTCAAGCA-----GAAATCACACAGTTTAACAAATATTG 3737
 QY 1239 GlnGluGluMetLysAsnIleThrIleGluArgAsnGluLeuGlnIleuThrAspPheGluAsp 1258
 Db 3738 CAAGATATGAACAATTACTTCAATTAAAGACAGAAAGAAAGAACCAAC----- 3788

QY 1259 LeuLysAlaGlnHisAspSerLeuLysGlnAspLeuSerGluLysIleGluGlnSerIle 1278
 Db 3789 ---AGCGAAGAAACAGAAATAATTGAAGAGAACTCTGCAACCTCTGTCACCCAG 3845
 QY 1279 GluThrGlnAspGluLeuArgAlaAlaGlnGluGluLeuArgGlnGluGlnLeuVal 1298
 Db 3846 AATCTGAAGACAGATCTTCAGAGAAAGAAAGATTATGCTGACCTGAAAGAAACTG 3905
 QY 1299 AspSerPheArgGlnGlnLeuLeuAspCysSerValGlyIleSerSerProAsnHisAsp 1318
 Db 3906 ACTGATGCCAAAACCCAGATT----- 3926
 QY 1319 AlaValAlaAsnGlnGluLysValSerLeuGlyGluValAsnSerLeuGlnSerLeuMet 1338
 Db 3926 ----- 3926
 QY 1339 LeuArgGlyGluArgAspGluLeuGlnIleuThrSerCysLysAlaLeuValSerGluLeuGlu 1358
 Db 3927 -----AAGCACTACAGAAAGAGATATCTGTAATGCTGATGAGATAA 3971
 QY 1359 LeuLeuArgAlaHisValLysSerValGluGlyGluAsnLeuGluIleThrLysLysLeu 1378
 Db 3972 TTACTGAGGATT-----AAAAT 3989
 QY 1379 AsnGlyLeuGluLysGluIleLeuGluLysSerGluGluSerGluValLeuLysSerMet 1398
 Db 3990 AATGAACCTGAGAAAAAAGAAAAACAGTGTTCAGGAATTAAGATTAAGAACAGCAAC 4049
 QY 1399 LeuGluAsnLeuLysGlu-----AspAsnAsnLysLeuLysGluGlnAlaGluGluTrp 1416
 Db 4050 ATTCAAGCAATCAAGAGAGAGATTAAATATCAAGAAAGTGAAGAACCTATACACAGTAT 4109
 QY 1417 SerSerLysGluAsnGlnPheSerLeuGlnValPheSerGlySerGlnLysLeuVal 1436
 Db 4110 GAGACAGCATCGAAAGATCTTAATGTTAAAG-----AAATTAAT 4151
 QY 1437 AspGluIleGluValLeuLysAlaGlnLeuLysAlaAlaGluGluArgLeuGluLys 1456
 Db 4152 GAAGACATG-----CGAATGACACTAGA 4175
 QY 1457 AspArgAspTrpPheGluLeuValGlnThrAlaAsnThrAsnLeuValGluLysLeu 1476
 Db 4176 GAACAG-----GAACAAACTCAGTAGAAGACAGATCAAGCTTCGACGCTAAATTA 4226
 QY 1477 GluThrProLeuGlnAlaAspHisGluLysAspSerIleAspArgSerLeuLysMet 1496
 Db 4227 GAG----- 4229
 QY 1497 GluIleLysValLeuGlyGluLysLeuGluArgAsnGlnTrpLeuLeuGluArgLeuGln 1516
 Db 4230 GAAGTTGAAGGCTGCCACAGAAATTTGAAAAATGGAAGAAATGCAATGATTTGGA 4289
 QY 1517 GluGluLysLeuGluLeuSerAsnLysLeuGluIleLeuGlnLysGluMetGluThrSer 1536
 Db 4290 ACCAAAACAAATCAAGGTCAAAATTA-----GAACATGAGAACACACACAGAT 4337
 QY 1537 ValLeu-----LeuLysAspAspLeuGlnGlnLysLeuGlnSerLeu----- 1550
 Db 4338 GTGCTTGAAGACCTCACTAATCTTCAAGATGCTGTTACAGAGGTCTGAACAGAAATATAT 4397
 QY 1551 -----LeuSerGluAsnIleIleLeu-----Lys 1558
 Db 4398 GCTGATGAAGAAATGTTGAGAAATAATGATGCTTATCACTCAACGCAAGCAAGACGA 4457
 QY 1559 GluAsnIle---AspThrThrLeuLysHisSerAspThrGlnAlaGlnLeuGlnLys 1577
 Db 4458 GAGATATACGAATAATTAAGAGATGTAATAATGCTGAGACAGGACGCTTTTATTAAG 4517
 QY 1578 ThrGlnGlnGluLeuGlnLeuAlaLysAsnLeuAlaIleAlaLysAspAsnPro 1597
 Db 4518 CAACGAGATGAATGAA----- 4535
 QY 1598 IleThrGlnGluLysGluThrSerAlaAspCysValHisProLeuGluGluLysIleLeu 1617

Db 2413 GATTGGATCACAAATATCCAAATTTCTAACCCAGCTA---GATTTTGCTTCTAAACGT 2469
Qy 1421 AsnGlnPhe-----SerLeuGluGlu 1427
Db 2470 TATGAAATGCTGCAGATTAATGTTGAAGATATGCTGCAGAAATATACATCATCTGCATGAG 2529
Qy 1428 -----ValPheSerGlySerGlnLysLeuValAspGluIleGluValLeuLys 1443
Db 2530 AGAAATACGAACCTCAGCCCAACTCAAAAGCAAGAACAGATTAATCAATACATGACT 2589
Qy 1444 AlaGlnLeuValAlaIleGluIleGluIleGluIleLysAspArgAspTyrPheGluLeu 1463
Db 2530 CAAGATTGAGAGAGCAATAGAGAGCTGCTGCAGAA----- 2631
Qy 1464 ValGlnThrAlaAsnThrAsnLeuValGluLysLeu-----GluThrProLeu 1480
Db 2632 GTAAAGCAGAAATTTGAAGAGAAAGAAATCTTAATTTGTCGAAGTTGCTT 2691
Qy 1481 GlnAlaAspHisGluGluAspSerIleAspArgArgSerGluGluMetGluIleLysVal 1500
Db 2692 TCTCAGCAAAAGAGCTCTTTGTTAGCTGACAAAGGGGGCAAACTTA----- 2739
Qy 1501 LeuGlyGluLysLeuGluArgAsnGlnTyrLeuLeuGluArgLeuGln---GluGluLys 1519
Db 2740 CTGCTAACTAATCTGCAAAACATTCAGGAATACTGAGCATCTGAAACAGAAACCAAA 2799
Qy 1520 LeuGluLeuSerAsnLysLeuGluIleLeuGlnLysGluMetGluThrSerValLeuLeu 1539
Db 2800 CAAAGCCTTAGTACCCAGATAGAAAACCTGGAAACATGAGATC----- 2841
Qy 1540 LysAspAspLeuGlnLysLeuGluLeuSerLeuLeuSerGluAsnIleIleLeuLysGlu 1559
Db 2842 ---TCTCATCTTAAGAGAGAGTTGCAAAATGAGGTGCAACAAAGCATACACTTCTAGA 2898
Qy 1560 AsnIleAspThrThrLeuLysHisSerAspThrGlnAlaGlnLeuGlnLysThrGln 1579
Db 2899 AATCTAGAT-----GTTCACCTTTTATAGATCAAAAG 2928
Qy 1580 GlnGluLeuGlnLeuAlaLysAsnLeuAlaIleAlaLysAspAsnLysProIleThr 1599
Db 2929 AGCAACTGTGATACAGAGCAAACTTT----- 2955
Qy 1600 GlnGluLysGluThrSerAlaAspCysValHisProLeuGluGluLysIleLeuLeuLeu 1619
Db 2956 -----CAT-----CTTAAC 2964
Qy 1620 ThrGlnGluLeuHisGlnLysThrAsnGluGlnGluLysLeuHisGluLysAsnGlu 1639
Db 2965 ACAAAAGAACTATTAAAA---AATGCTCAAAAGAAATATGCCACATTTGAACAGCAC 3018
Qy 1640 LeuGluGlnAlaGlnValGluLeuLysCysGlu---ValGluHisLeuMetLysSerMet 1658
Db 3019 CTGAGTAAATGTGAAGCCAAAGTTGCTTCTACAGTCTTCACAGAGAACTGTAAAGGTCAAG 3078
Qy 1659 IleGluSerLysSerSerLeuGluSerLeuGlnHisGluLysHisAspThrGluGlnGln 1678
Db 3079 CCTAGCAACAAAGAGATGTGATGATCTTGTAGTCACACTAAGACAGACAGAGAGCAG 3138
Qy 1679 LeuLeuAlaLeuLysGlnGlnMetGlnValIleThrGlnLysLysGluLeuGln--- 1697
Db 3139 GTGATGACTTTAAAGGAGAACTCAAAACAAAGACGAAATGTGAAACAAATTCAGAGA 3198
Qy 1698 -----GlnThrHisGlnHisLeuThrAlaGlu-----ValAspHisLeuLys 1711
Db 3199 ATGTTACTAGTTTAAAGAAATCCCTGAACAGAAACAGGTGACAGAGAGGCGT 3258
Qy 1712 GlnAsnIleGluLeuGlyLysAsnPheLysAsnGluAlaGlnGlnLysThrThrLysGlu 1731
Db 3259 AAGAAATTTGAAAGTTCGTTA---AAAGAGTCAGCTAATTTCAAGACAGCTTGGA 3312
Qy 1732 GlnCysLeuLeuAsnGluLysLeuGlnLysGluLeuGlnLysSerGlnHisArgLeuGlnCysGlu 1751
Db 3313 AAGAAATTTGATGAGATGAGAAAGAAACAGAACTTCAG----- 3354

Qy 1752 IleGluGluLeuMetLysSerLeuLysAspLysGluSerAlaIleLeuGluIleThrLeuLysGlu 1771
Db 3355 -----GATGATTAAGAAAGCCATGAGAGCATG----- 3384
Qy 1772 SerGlnLysValIleAsnLeuAsnGlnGluMetGluMetValMetLeuGluMetGlu 1791
Db 3385 ---GAACAAAGTTTATGATGAAGAAACACTTTCTAGTGTTCAGAAATGCAATCA 3441
Qy 1792 GlnLeuLysAsnSerGlnArgThrValIleAlaIleGluArgAspGluLeuGlnAspLeu 1811
Db 3442 GAAGCTTCTCAGAGCAACAGCAAGCTTAAAGATGAGCAGCAACCCAGAGCTACGT 3501
Qy 1812 ArgGluSerValGluMetSerIleGluThrGlnAspAspLeuArgLysAla----- 1828
Db 3502 CAGGAACAGCTAAATTAAGCTGTGAAAGCTCAGAAATAGAGAGAAATTTGATGCTG 3561
Qy 1829 -----GlnGluAlaLeuGlnGlnGlnLysAspLysValGlnLysLeuThrSer 1844
Db 3562 CATGCTGTGATGTGTGAAGCTTACAAAGCTCAGAAAGCAGAGCTGTCAAATGGCATCA 3621
Qy 1845 GlnIleSerValLeuGlnGlnLysIleSerLeuLeuGluAsnGlnMetLeuThrAsnVal 1864
Db 3622 GTCCGTGACATTTGGAGAAACACACAGAAAGCAGAAATCAAGTGTGGAGTGA 3681
Qy 1865 AlaThr-----ValLysGluThrLeuSerGlu 1873
Db 3682 GCATCTTGGGAGAAAGAGAGAAATGTTAAAGATGAAGTTCCAAATGCTATGCTGC 3741
Qy 1874 ArgAspAspLeuAsnGlnSerLysGlnHisLeuPheSerGluIleGluThrLeuSer 1892
Db 3742 TGTAAGATCTGGAGAAACAAACAGATTCTTCAATGATGATGCAAAATTAATGAG 3801
Qy 1893 -----LeuSerLeu----- 1895
Db 3802 AAGTCGTGCTGTGTGAAGAGAGTGTACAAAGTCCATGATGATATCTCAGTGA 3861
Qy 1896 -----LysGluLysGlu 1899
Db 3862 GAAGAAATCTCAAGAAACAAATTTTGGAAATTTCTCAGATTTATACGACAGAAAGAA 3921
Qy 1900 -----PheAlaLeuGluGlnAlaGlu----- 1906
Db 3922 ATTGCTGAATAGATTGAGGTGAGGTGCTGAGAGTGTGCTGATTCGAAAGGCTT 3981
Qy 1907 -----LysAspLysAlaAspAlaIleArgLysThr 1916
Db 3982 GAATCTTTAGAAAGAGAGCTTCAGAGCTCAGAACTGAAATGCTTAATGCTGAAGGAGAA 4041
Qy 1917 IleAspIleThrGluLysIleSerAsnIleGluGlnLeuLeuGlnGlnAlaThrAsn 1936
Db 4042 GTCCAGTAACTGCAAAACAAATGGCTCAGATGAGAACTGATGAAGAAACTGAACA 4101
Qy 1937 LeuLysGluThrLeuTyrGluArgGluSerLeuIleGlnCysLysGluGlnLeuAlaLeu 1956
Db 4102 ATGATGTAGTTATGAGAACCAATAAATGCTTAAGAGAGAGAGAGAGCTGAACAG 4161
Qy 1957 AsnThrGlnHisLeuArgIleThrLeuLysSerLysAspLeuAlaLeuGluLysMetGlu 1976
Db 4162 GATCTACAGAAATCAAGCAAGAGTGAAGAACTGAGATTAATTTTAACTTCA 4221
Qy 1977 GlnGluArgAspGluAlaAlaAsnLysValIleAlaLeuThrGluLysMetSerSerLeu 1996
Db 4222 GAAGCAATGCTGAGCTGAGTGAAGAAAGCGTATGTTGAGGAGCAAGAACTCTTA 4281
Qy 1997 GlnGluGlnIle-----AsnGluAsnValIleThrLeuLysGluGlu 2010
Db 4282 GAAGAGATGTCAAAGCTTGGAAAGCAGCTAACCGATCAAGATCTGATGAAG----- 4329
Qy 2011 GlnGlyGluLysGluThrPheTyrLeuGlnArgProSerLysGlnGlnSerSerGln 2030
Db 4330 -----CAACAGAAAGATCCAGAT 4347

OY 2031 MetGluLeuArgGluSerLeuLysThrLysAspLeuGlnLeuGluAlaGluLys 2050
 4348 ACAGAAAGATATCGAGACCCCTTCTGAAAGGAAGTTAT-----ACTAG 4395
 OY 2051 GluIleSerGluAlaThrAsnGluIleLysAsnLeuThrAlaLysIleSerSerLeuGlu 2070
 4396 CGATATCAACAATTCAGACAGAAATTCGTAGACTTAAGCTGAATTCGAAGATCAAT 4455
 OY 2071 GlnGluIleLeuGlnAsnAlaSerIleLeuAsnGluAlaValSerGluAlaGluAsnLeu 2090
 4456 GCATCTTACTATACACAGAACTTAATTCAGAGTCTGAAGGAAGTCTAAATTAAGTA 4515
 OY 2091 ArgHisSerLysGlnLeuValSerGluLeuGlnLeuSerLeuThrLeuLysSer 2110
 4516 AGAATCGAAAGAAACCAATTCAGAGAGACTTGATGCCAAATTAATTCATCCAGAA 4575
 OY 2111 ArgAspHisAlaPheAlaGlnSerLysArgGluLysAspGluAlaValAsnLysIleAla 2130
 4576 AAGATCAAAACTATATCTCAAGTTAAGAAATTCAGAGTACAAAGACTCAATATGAA 4635
 OY 2131 SerLeu---AlaGlnGluIleLysIleLeuThrLysGluMetAspGluPheArgAspSer 2149
 4636 GAATCTTAAGACACACAGATTAAGTTATGGACATCGGCTCAGTCTCTGGAGACCAT 4695
 OY 2150 LysGlu-----SerLeuGlnGlnIleSerSerHisLeuSerGluLeuLysThr 2166
 4696 CAGGACACACATGTTCCAGTCCGAGAAATG---CAGGAATCTCAAGAAACGCTCAACCA 4752
 OY 2167 TyrLysThrGluLeuGlnMetLeuLysGlnGlnLysGluAspIleAsnAsnLysLeuAla 2186
 4753 GGTGAACAAACAAATCAAACTCTGAAAGTCAAGTACAGATCTGCAAGAACATTAATCT 4812
 OY 2187 GluLysValLysGluValAlaAspGluLeuGlnHisLeuSerSerLeuLysGlnLeu 2206
 4813 GAAAGAAAGACACAGACCAAAATCTCCAGAAACAGACTGTGCACTTCACTGTAACCTT 4872
 OY 2207 AspGlnIleGlnMetGluLeuArgAsnGlnLysLeuArgAsnTyrGluLeuLysGluLys 2226
 4873 TCACGACTTCGTCAAGATTCATCAAGATAGAACACACAGAGAGAGACAGCTC-----CGA 4926
 OY 2227 MetAspIleMetGluLysGlu----- 2233
 4927 CAACAGATATCTGAAGAGAGAAACCAACGAAAGCTATTGTACACAGCAAAAGTCAAA 4986
 OY 2234 IleSerValLeuArgLeuMetGlnAsnGlnProGlnGlnGlnLysAspValAlaGlu 2253
 4987 ATTCGACACTTACCTGCTGTAAAGATCACTTAAGAAATGAGAGCTTAACAA 5046
 OY 2254 ArgMetAspIleLeuGlnSerArgAsnGlnIleGlnLeuMetGluLysIleSer 2273
 5047 AGGAATGAGCCTTACAT-----CAGCAAAAGATCAATTCGATGTCATTAAT 5097
 OY 2274 AlaValTyrSerGlnHisThrLeuLeuSerSerLeuSerGluLeuGlnLysGlu 2293
 5098 GGCCTTAAGTCCCAATATGAAGTCAAGTTCGCTTGAAGAAAGCAACTCAGAGGAGCAT 5157
 OY 2294 ThrGluAlaHisLysHisCysMetLeuAsnIleLysGlu-SerLeuSerSerThrLeu 2313
 5158 CAAGAGAGACAC-----CTTGACACAGAGATGAGCTCAAGAACCTTCTTAAT 5205
 OY 2313 ArgSerPheGlySerLeuGlnThrGlnHisValLysLeuAsnThrGlnLeuGlnThrLe 2333
 5206 AAGTCCCGTGAACAGACAGACAGATCATTTGAAA-----CAACTCCAGCTTC 5256
 OY 2333 LeuAsnLysPheLysValValTyrArgThrAlaAlaValLysGluAspHisSerLeuI 2353
 5257 G-----GTCAAGAGAC----- 5266
 OY 2353 LysAspTyrGluLysAspLeuAlaGlnGlnLysArgHisAspGluLeuArgLeuGln 2373
 5267 -----GAATTCGACACATCATCAGACCCACCAACAGCAATATCAAGCCA 5310
 OY 2373 n---LeuGlnCysLeuGlnHis-----GlyArgLysTr 2384

Db 5311 ACTGCTGTGTGCTACTCCAAAT-AAAAGACAGCTGCAGCTATGGCTGGAAATAATGTC 5369
 OY 2384 pSerAspSerAlaSerGluLeuLysPheCysGluIleGluPheLeuAsnGluLeu 2404
 5370 AACACCCAGGCGTATTCGCCCAATGTTTACACCTGCACACTGTTTCAAAATCCACTTAC 5429
 OY 2404 uphLeuLysAlaAsnIleLeuGlnSerValGlnAspAspPheSerLeuValGlnValPhe 2424
 5430 TACCCCAACAGCTACAGTGAATGCCCTACACAAAGTGA----- 5469
 OY 2424 LeuAsnGlnValGlySerThrLeuGlnGlnLeuGlnHisLysLysGlyPheMetGlu 2444
 5470 -----TCACAGAAAGCTATGCAGTCAACAGAGGCGCTGCGA 5504
 Db 2444 nTTPLeuGlnGluPheGlyAspLeuHisValAspAlaLysLysLeuSerGluMetGlu 2464
 5505 ACATGTTCCAGTTTTCGAAAGCACACACTGGATCCGTTCTCTACTAGTCTTAATGTCCA 5564
 OY 2464 nGlnGluAsnArgIleAlaSerThrIleGlnLeuLeuThrLysArgLeuLysAlaVal 2484
 5565 GCCTTCATCTCAACCTATTCTTAACTGTTCACCAACAAACA---CAGCTACAGCTTT 5621
 OY 2484 ValGlnSerLysIleGlnArgGluIleThrValTyrLeuAsnGlnPheGluAlaLysLe 2504
 5622 TGTGCAACCCACATCAACAGATCAT-----CTTCAGATTCAGCCTGCGCA 5666
 OY 2504 uGlnGluLysLysGlnHisnLysGluLeuMetArgArgMetGlnHisHisGlyProse 2524
 5667 TCAGAGATTAATCTTCAAAACATAGTAGAGGTTCACAGTTCACACACTTACGCGGCTTC 5726
 OY 2524 ValSerValMetGluGlnGluAsnAlaArgLeuGluIleLeuLysThrValGlnAs 2544
 5727 TACTTCC-----ACAGCAGATTTTGGACACAGCTTCGGCTACCCCCAG 5768
 OY 2544 pGluSerLysLysLeuGlnSerArgIleLysMetLeuGluAsnGluLeuAsnLeuValLys 2564
 5769 TCTCTTCTTCCAAAGCGTACAGT-----GAAGAGGA----- 5802
 OY 2564 sAspAspAlaMetHisLysGluLysValAlaIleLeuGlnAspLysLeuSerArg 2584
 5803 -----GAGGATGACACCATAGAAC 5822
 OY 2584 gAsnAlaGluAlaGluLeuAsnAlaMetGlnValLysLeuThrLysLysGlnAsnLeu 2604
 5823 ATCAGACCAAGTCTGTATGATATACAGTGAATGCCCTTCCAAAGAG----- 5871
 OY 2604 uGlnAlaAlaMetLysGluIleGluAsnLeuGlnLysMetValAlaLysGluAlaValPhe 2624
 5872 -----TTGAAAAGTGCACACC 5888
 OY 2624 cTyrLysGlnGluIleAspAsnLeuLysThrLysValValLysIleGlnMetGluLysIle 2644
 5889 TGTAGCAATCGAGAGAAAGATTTATGGCAGAAAGAAAGTATGATGAGAGGTAGAGACTCA 5948
 OY 2644 eLysTyrSerLysAlaThrAspGlnGluIleAla-----TyrLe 2657
 5949 GGTATACACACAGATTTCTAAGATTCATTTGGAGAAAGAGATTACCCAGGAGATTAAT 6006
 OY 2657 uLysSerCysLeuGluLysAspLysGlnGluLysLeuArgLysGluLysGluLysLeuArg 2677
 6007 -----ACACCTATGGAAGACAGTGAAGAAACCTCTCACTCAATATAGATCTTGCGCC 6062
 Db 2677 gAlaGlnAlaAspAsnAspThrThr----- 2685
 OY 2686 ---ValCysValProLysAspTyrGlnLysAla----- 2695
 Db 6123 TGTCAATTTGATTTGACAGTATGATGAAGAGAGATGAGAAAGATGATGATGATCA 6182
 OY 2695 ----- 2695


```

Db 6183 AGATGACACAGGATGGAGATGAGGTGAAGATAGTAAGAAGAACTGGTAGTGGCA 6242
QY 2696 -----SerThrpheProValTh 2701
Db 6243 TGGCAATGATGATTGACACTGATGCTGAGCGGTGGATGAGGACCTGATCCAGTAC 6302
QY 2701 T-----CysGlyGlySerGlyIleValGlnSerThrAlaMetLeuVa 2716
Db 6303 AGAAGACAGAAAGATGATGGGTGAGGTGAAGT----- 6336
QY 2716 IleuGlnSerGluLysAlaIleuGluArgGluLeuSerHisTyrLysLysTyrH 2736
Db 6337 -----AsnCAGAGAGCTGCTGATCTCTCAAAACAGTGTGAAGAAATACAGGTCTGC 6389
QY 2736 ShISLeuSerArgThrMetSerSerSerGluAspArgLysThrLysAlaLysSerAs 2756
Db 6390 AGAATCTCTTTCTTCAGAGAGTTCTAGAGAACAAAGCCCTCATACAGCA---TCTGA 6446
QY 2756 pAlaHisSerSerHisThrGlySerSerHisArgGlySerProHisLysThrGluThrTy 2776
Db 6447 AAGACAGGCGCCCTGAGCACTTCACCGACGAGCCGCCACACAT----- 6492
QY 2776 rArgHisGlyProValThrProGluArgSer---GluMetProSerLeuHisLeuGlySe 2795
Db 6493 -----CCACTCCCGCCCAAGACTGACCAATTCATGCCCACTCAGAGAGTGGACC 6542
QY 2795 rProLysLysSerGluSerSerThrLysArgValLysSerProAsnArgSerGluIleTy 2815
Db 6543 ACCA-----GTTCCAGAGAAATTCAGATGACGCCAGCCAGCTCGTAGG 6584
QY 2815 rSerGlnLeuValMetSerProGlyLysThrGlyMetHisLysHisIleLeuSer----- 2833
Db 6585 AGCTGGCCCTTCAGTGTACTCCAGAAATGAGTGGCAGTCAACAGCAATTTTGTGATGATGA 6644
QY 2834 -----ProSerLysValGlyLeuHisLysLysArgAlaLeuSerProAsnArg 2849
Db 6645 AGACAGACAGTTCACAAAGTACTCAACTCTT-----GAGGTGCCACATCG 6689
QY 2849 gSerGluMetProThrGlnHisValIleSerProGlyLysThrGlyLeuHisLys----- 2867
Db 6690 TACTATGATGATTTGCTGAAGCAATTCATTCGCCGAGGTGTGCTGTGTCCTAGATTCGG 6749
QY 2868 -----AsnLeuThrGluSerThrLeuPheAspAsnLeuSerSerProCy 2882
Db 6750 GTTTGGCCACCTGAAAGATATGCCACAAACAAGT-----TCTATCTACTCTGATCT 6800
QY 2882 sLysGlnGlnLysValGlnGlnAsnLeuAsnSerProLysGlyLysLeuPhe----- 2899
Db 6801 TGGCAGCTTCTCTCTCAAGAGGTTTAGAATGATATGAAACACCCCTGTCTCTAGCTCA 6860
QY 2900 -----AspValLysSerLysSerMetProTyrCysProSerGln----- 2912
Db 6861 TGAAGAAAGATCAGGTGGCCGAAGTGTCCACACTACTCATACTAAGATAGCAGCCCACT 6920
QY 2913 -----PhePheAspAsnSerLysLeuGlyAspPheSerGlu 2924
Db 6921 GACTGTATTTACTGAGACACCACCTCTGATGCTTCGAA 6960

```

```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: Incyte ID No. US20020137081A1 2171870CB1
US-10-044-090-359

Alignment Scores:
Pred. No.: 4.98e-48 Length: 7792
Score: 891.50 Matches: 540
Percent Similarity: 37.15% Conservative: 446
Best Local Similarity: 20.35% Mismatches: 959
Query Match: 6.04% Indels: 710
DB: 12 Gaps: 97

US-09-150-867-1 (1-2954) x US-10-044-090-359 (1-7792)
QY 600 SerArgGluHisSerIleAsnAlaGluValGlnThrAspValGluLysGluValAlaArg 619
Db 111 TCCCGCCCGGAGTCTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 170
QY 620 LysGluMetSerValLeuGlyAspSerGly-----TyrAsnAlaSerAsnSer 635
Db 171 GCGGCGCGCCAGCGCTGTGGTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 230
QY 636 AspLeuGlnAspSerSerValAspGlyLysArg---LeuSerSerSerHisAspGluCys 654
Db 231 TCGGCGCGTGCACATCGGGGCTCCCTGTGCGCCACCCACATGCGGCGGCTGTGCGACCAAGTC 290
QY 655 IleGluHisArgLysMet-----LeuGluGlnLys 664
Db 291 CTGAGGCGCCAGCGAGCTGACACACAGCTGCCCAAGTCTGTCCAGAACAACTGAAAGATTC 350
QY 665 IleValAspLeuGluGluPheIleGluAsnLeuAsnLysLysSerGluAsnAspLysGln 684
Db 351 CTGCTGATTCAGCAATCCGAGATGATGCTGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 410
QY 685 LysSerSerGluGlnAspPheMetGluSerIleGlnLeuGluAlaIleMetAlaGlu 704
Db 411 GAGAGCGCAACAGTATTT----- 431
QY 705 LysAlaAsnAlaLeuGluGluLeuAlaLeuMetArgAspAsnPheAspAsnIleIleLeu 724
Db 432 -----GAATAGAAAGAGGTGTCTCCACAGTACGAGAGACTGTGAATGAACCCGA 485
QY 725 GluAsnGluThrLeuLysArgGluIleAlaAspLeuLysSerLysGluAsnGln 744
Db 486 GAGTGTCAAGCTTGGCGGCTGAGCTAGAGAACTCAACATCACTGAAGGACACTACT 545
QY 745 GluThrAsn---GluPheGluIleLeuGlu----- 753
Db 546 GAGAAACAAAGAACTTGAATTCACAGATCCGCAATATTCATTCAGACCAATTT 605
QY 754 LysGluThrGlnLysGluHisGluAlaGlnIleuIleHisGluIleGlySerLeuLysLys 773
Db 606 ACAAGAACAAAGAGAAATTAAGAACTGAG-----AAAAGAC 644
QY 774 LeuValGluAsnAlaGluMetTyrAsnGlnAsnLeuGlu-----GluAspLeuGlu 790
Db 645 TTAATTAGAACCAATGAGACTATCTCAAGACTTGAATTAATTACAGAGGATGTATAA 704
QY 791 ThrLysThrLysLeuLysGluGlnGluIleGlnLeuAlaGluLeuArgLysArgAla 810
Db 705 CGTCTGAATGAAACCTTAAGAAACCAATACAAACAAGGCGTCACTTGAATTAATTAATTTG 764
QY 811 AspAsnLeuGlnLysLysValArgAsnPheAspLeuSerValSerMetGlyAspSerGlu 830
Db 765 GATGAACCTTCAAGCTCT-----GAGTTTCTGTTAAGATTCAGAA 806
QY 831 LysLeuGluGluGlu-----IlePheGln 838
Db 807 AAACGCTTGAGCAGCAAGAAAGAAATTCATACATAGTCAAGAAATACATGCTGAATACAGAG 866
QY 839 LeuLysGlnSerLeuSerAspAlaGluAlaValThrArgAspAlaGlnLysLysSer 858

```

```

Db 867 TTGAACCAAACTGATGACTCTGCTCTTGGACAGAGAAAAGGAAATGAGATT----- 923
Oy 869 PheLeuArgSerGluAsnLeuGluLeuGluMetGluAspThrSerAsnTrpIyr 878
Db 924 -----CTAGAGCTTAAATGTAATCTTGA----- 947
Oy 879 AsnGluLeuGluAlaAlaSerLeuPheGluGluGluGluThrGluLeuSerAsn 898
Db 948 AATAAAGAGAGAGAGGTTCT-----AGACTGGAGAACAAATGAAATGGCG 992
Oy 899 TyrIleuMetGluAlaAspLeuGluGluGluGluGluGluGluGluGluGlu 918
Db 993 TTAATAAATCAATGACATCTTCAAAAGCATGTGAG----- 1031
Oy 919 TyrLeuAsnGluLeuLeuAlaGlyValProArgAspLeuLeuSerArgValGluLeu 938
Db 1032 -----GATCTGTGTGACCAAAATTA----- 1049
Oy 939 GluIleuValSerGluPheSerGluLeuGluGluAlaLeuGluGluGlu----- 956
Db 1050 -----AAAGAGCCAAAGACACAGCCAGTATGGAGAGAAATTCAC 1094
Oy 957 AsnAlaLeuGluAsnGluValThrCysLeuSerGluTyrIleuPheLeuProAsnGluVal 976
Db 1095 AATGATTAATGATGCCACATATAAATCTTCTAATTTGTACAAAGATGCCGATGATGCTCA 1154
Oy 977 GluCysLeuIleuAsnGluIleuSerGluIleuGluGluGluGluGluGluGlu 995
Db 1155 GAAGCAAGAGCAATGAACCTACCGGACGTAGAGAACTACACAAATTTTGAAGAA 1214
Oy 996 GluGluGlu-----HisSerAlaSerIleuLeuSerGluGlu 1008
Db 1215 GCTGGTGAAGCCAAAGCAATACAGATCATCTTCTAGAGAGTGGAGAAATCCAAAGAT 1274
Oy 1009 IleuIleuGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 1028
Db 1275 CAAATGGAAAAAAGAAATGCTTGAGAAATAGGAGATTGGAGAGAAATTAGAAGATGCA 1334
Oy 1029 GlnSerIleuValGluGluGluGluGluGluGluGluGluGluGluGluGlu 1040
Db 1335 AATGACCTTCTTGTGCAAAAGCAAGGACCAATATGTCGAAAGAGAGCTTGGCC 1394
Oy 1041 GluMetIleu-----LysMetHis 1046
Db 1395 GCCATGCTCTCTACTGACGACGACCTAGTACATAGTAAACCTGGAGTAACTA----- 1451
Oy 1047 AspAspLeuPheGluIleuGluGluGluGluGluGluGluGluGluGluGluGlu 1066
Db 1452 ACTGAGCTCTAATGCTTATGTGAAACTCAGATCACTTG-----CTTTGGAGAA 1505
Oy 1067 MetGluAsn-----LeuIleuGluGluGluGluGluGluGluGluGluGluGlu 1080
Db 1506 CTAGAGAAACAAAGAAATTAATAGTACATAGTAAATGAAATGAAAGAAATGGAAGCCAA 1565
Oy 1081 IleuAlaSerThrIleuGluGluGluGluGluGluGluGluGluGluGluGluGlu 1100
Db 1566 GCACCAATTTTGAAGGCGGAGGAGTGAATGAAACGTCGACGAGAAACCTGATGACAACT 1625
Oy 1101 GluIleuGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 1120
Db 1626 TTATCTGTTAAGCTGACAAAGCTATGAAAGAGATTGACGATGACAGAGAGACACTGAT 1685
Oy 1121 SerLeuProPheSerLeuValGluGluGluGluGluGluGluGluGluGluGlu 1140
Db 1686 -----AAAGCCAAAGCAATCATCTGTA----- 1709
Oy 1141 TyrHisAsnLeuIleuAlaLeuAlaThrGluArgAsnAsnIleuMetValCysLeuGluThr 1160
Db 1710 -----CTTGAGACA 1718
Oy 1161 GluArgAsnSerLeuGluGluGluGluGluGluGluGluGluGluGluGluGlu 1180

```

```

Db 1719 GATTAATCGAAGATGGAATACAAATGTAAGATCTTTCACAAACAGATTAAGAGCTT----- 1775
Oy 1181 AlaGlnSerIleuGluGluSerAspLeuGluGluGluGluGluGluGluGluGluGlu 1200
Db 1776 TTGATGGAACCTGMAAACCAAGGGTAAACCGTAATTCGGATGAGAGATGAACTCT 1835
Oy 1201 ValIleuLeuLeuGluGluMetGluLeuLeuGluGluGluGluGluGluGluGluGlu 1220
Db 1836 GCTGATATGATAGTATCATCTGAGATATATCATCAGACATCTAGTA-----TCTTACAAAT 1892
Oy 1221 IleuGluLeuGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 1240
Db 1893 ATTGAAGACCTTCACAAACAAATCAACGCTCTTGTAGTGGCTTGGAGAGCTTGGGAA 1952
Oy 1241 GluMetIleuAsnIleuThrIleuGluArgAsnGluGluGluGluGluGluGluGlu 1260
Db 1953 -----ACCAAGAAAGAGAAAGAAACAAACAACT----- 1982
Oy 1261 AlaGluHisAspSerLeuGluGluAspLeuSerGluAsnIleuGluGlnSerIleuGluThr 1280
Db 1983 -----TCATCCAAATTCAGCTGATGCTGACGCTCAAACTTGAAGTGCCTTACTGAA 2033
Oy 1281 GlnAspGluLeuArgAlaAlaGluGluGluGluGluGluGluGluGluGluGluGlu 1300
Db 2034 CTAGAACAACTCCGCAATCAGACAG-----CATCAATTCGACGCTTGTGATGCC 2084
Oy 1301 PheArgGluGln-----LeuLeuAspCysSerValGlyIleuSer 1314
Db 2085 ATAGTTCGTCAGGCTGATATGATACCGTATTTATTTTCACAAACACAGAGATGGCAAT 2144
Oy 1315 ProAsnHisAspAlaValAlaAsnGluGluGluGluGluGluGluGluGluGluGluGlu 1334
Db 2145 CCATTACATGCTTCA-----AGCTTACATATGCTT-----TCTCTT 2180
Oy 1335 GlnSerGluMetLeuArgGluGluArgGluGluGluGluGluGluGluGluGluGlu 1349
Db 2181 GCATCAATCCAAACAGTCAATGATCATCAAGACTGTTTCCACTGCTCCAGTACCT 2240
Oy 1350 -----CysIleuAlaLeuValSerGluGluGluGluGluGlu 1360
Db 2241 GTTATGAAATCAACAGAGCTATAGAGCTTACAGGCTGCTTAAACAGTTGACGAAAT 2300
Oy 1361 ArgAlaHisValIleuSerValGluGluGluGluGluGluGluGluGluGluGluGlu 1380
Db 2301 TTGAGAACTACAAAAAAGCAAGCAAAATGAAAAATA----- 2342
Oy 1381 LeuGluGluGluIleuGluGluGluGluGluGluGluGluGluGluGluGluGlu 1400
Db 2343 -----CAAAATGAGCAGCTTGAGAAACCTTCACAAACAAAGTTACA 2381
Oy 1401 AsnLeuGluGluAspAsnAsnIleuGluGluGluGluGluGluGluGluGluGluGlu 1420
Db 2382 GATTTGCGATGCAACAAATATCCAAATTTCTACCCAGCTA-----GATTTGCTTAAACCT 2438
Oy 1421 AsnGlnPhe-----SerLeuGluGlu 1427
Db 2439 TATGAAATCTCGAATATGTTGAAGATATCTCGAGAAATATACATCACTTCAATGAG 2498
Oy 1428 -----ValPheSerGluGluGluGluGluGluGluGluGluGluGluGlu 1443
Db 2499 AGAAATCAGAAATCTCACTCCACAACTCAAAAGCAAGAACAGATTTCATATGAGATGACT 2558
Oy 1444 AlaGlnLeuLeuAlaAlaGluGluGluGluGluGluGluGluGluGluGluGluGlu 1463
Db 2559 CAAGATTTGAGAGACCAAAATGAGAGCTGCTGCGAGAA----- 2600
Oy 1464 ValGlnThrAlaAsnThrAsnLeuValGluGluGluGluGluGluGluGluGluGluGlu 1480
Db 2601 GTTAGAGCAAGAAATTTGAAGAAAGAAAGAAATGCTTAAATTTGCTGAGATGCTGCTT 2660
Oy 1481 GlnAlaAspHisGluGluAspSerIleuAspArgArgSerGluGluGluGluGluGluVal 1500
Db 2661 TCTCAGCAAAAGAGCTTGTAGCTGAACAAAGGGGCAAAACCTTA----- 2708

```

QY 1845 GlnIleSerValLeuGlnGluIlyrIleSerLeuLeuGlnIleValSerIleuValAspVal 1864
 Db 3531 GTCCGTCAGCATTTGGAAACAAACACACAGAAAGCAAAATTCACAGTTGTTCAGATCGTAA 3550
 QY 1865 AlaThr-----ValIySgluThrLeuSerGlu 1873
 Db 3651 GCATCTTGGAGGAGAAACAGACAGAAATGTTAAAGATGATGATTCCTCAATATGCTGTCTGCC 3710
 QY 1874 ArgAspPheLeuAangInSerIySgInIhSLeuPheSerGluIleGluThrLeuSer--- 1892
 Db 3711 TGTGAAGATCTCGAAGAAACAAACACAGATATTCTTCATGATGATGATGCAAAATTAAGTAC 3770
 QY 1893 -----LeuSerLeu----- 1895
 Db 3771 AAGGTCGTGGCTCTGTGAAGAAAGGTGTACAAAGTCCACTGAATGTATCTCTCAGTGA 3830
 QY 1896 -----LysGluIySglu 1899
 Db 3831 GAAGGAAATCTCAAGAACAAATTTGGAAATTCAGATTATTCAGACGAGAAAAAGAA 3890
 QY 1900 -----PheAlaLeuGluIleValGlu----- 1906
 Db 3891 ATTGCTGAATAGCTTTGAGTGGCTCAGGTTGAGAGTCCGCTATTCGACAAAGGTT 3950
 QY 1907 -----LysAspIhSaAspAlaIleArgIyThr 1916
 Db 3951 GAACTTTAAAGAAAGAGCTCGACGAACTCGAAGATGACTAAATGCTGAAGGAGAA 4010
 QY 1917 IleAspIleThrGluIlyrIleSerAnIleGluGlnLeuLeuGlnIleValIleThrAsn 1936
 Db 4011 GTCCAGGTAATCTCAAAACAAATGGCTCAGATGAAAGCAATGAAAGAAATGAAACA 4070
 QY 1937 LeuIySgluThrLeuIyThrGluIyGluSerIleGlnIySgInIleValIleValIleu 1956
 Db 4071 ATGAATGACTAGTTAGGAGACCAATAATATCTTAAGAGAGAGAGAGAGAGCTGAACAG 4130
 QY 1957 AsnThrGlnIhSLeuArgIyThrLeuIySerIyAspLeuAlaLeuGluIyLysMetGlu 1976
 Db 4131 GATTCACGCAATCGACAGAAAGAGAGAACTGAACTGATGATTTACCTTCACA 4190
 QY 1977 GlnGluIyArgSgluAlaIleAsnIyValIleAlaLeuThrGluIyMetSerSerLeu 1996
 Db 4191 GAAGCAATGCTGAGCTGAGTGAAGAAACGGTATTTGACAGCAGAGAAAGCTCTTA 4250
 QY 1997 GlnGluGlnIle-----AsnGluAsnValIleThrLeuIySgluGly 2010
 Db 4251 GAAGAGAGATGCAAAACGTGGAAGACGTAACACACTGTAGT----- 4298
 QY 2011 GlnGluIyIySgluThrPheIyThrLeuGlnArgProSerIySgInIleSerSerGln 2030
 Db 4299 -----CAspGluAsnArgTCCAGT 4316
 QY 2031 MetGlnIleuArgIySerLeuIyThrIyAspLeuGlnIleuGlnGluIleValIySglu 2050
 Db 4317 ACAGAGATATTCGAGAGCTCTTTCGAAAGAAAGTTCAT-----ACTAAG 4364
 QY 2051 GlnIleSerGlnAlaThrAsnGluIleIyAsnIleuThrAlaIySgluSerSerLeuGln 2070
 Db 4365 CGTATTCACAAATGACAGAGAAATTTGTAGACTTAAAGCTGAATTCGAACATCAAT 4424
 QY 2071 GlnGluIleLeuGlnAsnIleSerIleIyLeuAsnGluIleValSerGluArgIleuAsn 2090
 Db 4425 GCATCTTGTACTACACAGAACTTAATTCAGAGTCTGAAGAGAACTTAAATTAAGTA 4484
 QY 2091 ArgIhSerIySgInIleuValSerGluLeuGlnIleuSerLeuIleuSerIleuIySer 2110
 Db 4485 AGAAGTGAAGAAACCAACCTCGAAGAGCTTAAGTGCACAAATATGATATTCACAGAA 4544
 QY 2111 ArgAspIhSaIlePheAlaGlnSerIyAspGluIySgIyAspGluIleValAsnIySglu 2130
 Db 4545 AAAGTCAAACTATTACTCAAGTTAAGAAATTTGGACGAGGTACAAAGCTCAATATGAA 4604
 QY 2131 SerLeu---AlaGlnGluIleIySgluIleuThrIySgIleuMetAspGluPheArgAspSer 2149

Db 4605 GAAGTTAAAGACACAGATAGGTATGAGACATGGGCTCACTCCCTGGAGACCAT 4664
 Qy 2150 Lysglu-----SerleuglnlgluInserHisleuSerGluLeucysThr 2166
 Db 4665 CAGAGCAGCATGTTCTCCAGGAAATG---CAGAACTCAAGAAACGCTCAACCA 4721
 Qy 2167 TyrLysThrGluLeuGlnMetleuLysGlnGlnLysAspIleAsnAsnLysLeuLa 2186
 Db 4722 GCTGAACCAAAATCAAAATCTGTAAGTCAAGTGAATCTGCACAGACATTTACT 4781
 Qy 2187 GluLysValLysGluValAspLysLeuGlnHisleuSerSerleuLysGlnLeu 2206
 Db 4782 GAAAAGAGACAGAGAAATCTCCAGAGACAGACTGTGCACACTTCAGCTCAACTT 4841
 Qy 2207 AspelInleGlnMetGluLeuArgAsnGlnLysLeuArgAsnTyrGluLeuLysGln 2226
 Db 4842 TCACGACTTCCTCAGAGACTTCAGATGAAACACACAGAGAGACACTC---CGA 4895
 Qy 2227 MetAspIleMetGluLysGlu----- 2233
 Db 4896 CAACAGATTAAGTAAAGGAAAGAAACCAAGAAAGCTATTGTACAGCAAAAGTCAAA 4955
 Qy 2234 IleserValleuArgleuMetGlnAsnGluProGlnGlnGlnLysAspValAlaGlu 2253
 Db 4956 ATTCACACTTAAGCTGCTGTAAGATCAGCTAAAGAAATGAGAGACTTAAACA 5015
 Qy 2254 ArgMetAspIleleuGlnSerArgAsnGlnLysGlnLeuMetGlnLysIleser 2273
 Db 5016 AGGATGAGAGCTTAGAT---CAGCAAGAAAGATGATGATGCTTCCGCTTTCT 5066
 Qy 2274 AlaValTyrSerGlnGlnHisThrleuLeuSerSerleuSerGlnLeuGlnLysGlu 2293
 Db 5067 GCGTAAAGTCCCAATATGAAGCTCGATTTAGTGGCTTGAAGAGAACTCAGGAGCAT 5126
 Qy 2294 ThrGluAlaHisLysHisCysMetleuAsnIleLysGlu---SerleuSerSerThrLeuSe 2313
 Db 5127 CAACAGAGAGAC---CTTGACAGAGAGATGAGCCCAAGCAACCTTCTAT 5174
 Qy 2313 ArgSerPheGlySerleuGlnThrGlnHisValLysleuAsnThrGlnGlnIle 2333
 Db 5175 AAGGTCCCTGAACAGCAGACAGATCAGATGAAA---CAAGTCCAGCTTCT 5225
 Qy 2333 uLeuAsnLysPheLysValValTyrArgThrAlaValLysGluAsnHisSerleuI 2353
 Db 5226 G-----CTGAAGAG----- 5235
 Qy 2353 eLysAspTyrGluLysAspLeuAlaGlnLysArgHisAspGluLeuArgleuG 2373
 Db 5236 -----GAAATGCCAGACATCAGACCCACCAACAGCCAAATATCAAGCA 5279
 Qy 2373 n-----LeuGlnCysleuGlnGlnHis-----GlyArgLysTr 2384
 Db 5280 ACTCTCTGTGTCTACTCCAGT---AAAGTACAGCTGCACCTATGGCTGCAAAATATC 5338
 Qy 2384 pSerAspSerAlaSerGlnGlnLeuLysPheCysGluIleGluPheLeuAsnGlnLeuLe 2404
 Db 5339 AACACCCAGGCTAGTATCCGCCCAATGGTTACACCTGACCTGTACAAATCCACATAC 5398
 Qy 2404 uPheLysLysAlaAsnIleIleGlnSerValGlnAspAspPheSerGluValGlnValPh 2424
 Db 5399 TACCCCAACAGCTACAGTATGATGACCACTACACAAAGTGAA----- 5438
 Qy 2424 eLeuAsnGlnValGlySerThrleuGlnGlnLysGlnHisLysLysGlyPheMetG 2444
 Db 5439 -----TCACAGAGAGCTATGACAGTCAAGAGGCGCTGTGGA 5473
 Qy 2444 nTyrLeuGlnGluPheGlyAspLeuHisValAlaLysLysLeuSerGlnLysMetG 2464
 Db 5474 ACATGTTCCAGTTTTGGAAACACAAAGTGAATCCGTTCTACTAGTCTTAATGTCCA 5533
 Qy 2464 nGlnGlnAsnArgArgIleAlaSerThrIleGlnLeuLeuThrLysArgLeuLysAlaVal 2484

Db 5534 GCGTTCTATCTCAACCTATTTTAACTGTTTCAGCAACAACA---CAGGCTACAGCTTT 5590
 Qy 2484 lValGlnSerLysIleGlnArgGluIleThrValLysleuAsnGlnPheGluAlaLysLe 2504
 Db 5591 TGTGCACCCACTCAACAGAGTCAAT-----CCTCAGATTTGACCTGCCCA 5635
 Qy 2504 uGlnGlnLysLysGlnGlnAsnLysGluLeuMetArgArgMetGlnHisIleGlyProSe 2524
 Db 5636 TCAGAGTTATGTTCAACATAGTAGAGTTTTCAGAGTTTCCAGATTCAGAGCTGCGCTTC 5695
 Qy 2524 ValSerValMetGlnGlnGlnAsnAlaArgleuGlnIleLeuLysThrValGlnAs 2544
 Db 5696 TACTTCC-----ACAGCACTATTGTGCACAGTTCGGCTACCCCGAG 5737
 Qy 2544 pGluSerLysLysleuGlnSerArgIleLysMetleuGlnAsnGlnLysleuValLys 2564
 Db 5738 TTCTCTTTCGCCAAGCGTACAGT-----GAGAGGAA----- 5771
 Qy 2564 sAspAspAlaMetHisLysGlyGluValAlaIleleuGlnAspLysleuLeuSerAr 2584
 Db 5772 -----GAGATACGACCATGAAAC 5791
 Qy 2584 gAsnAlaGluAlaGluLeuAsnAlaMetGlnValLysleuThrLysGlnAspAsnLe 2604
 Db 5792 ATCAGACCAAGTCTGTGATGATACAGTGGAAATGGCTCTCCAAAGAG----- 5840
 Qy 2604 uGlnAlaAlaMetLysGluIleGlnAsnLeuGlnLysMetValAlaLysGlyAlaValPr 2624
 Db 5841 -----TTGAAAAGTGCACACC 5857
 Qy 2624 cTyrLysGlnGluIleAspAsnLeuLysThrLysValValLysIleGlnMetGluLys 2644
 Db 5858 TGTAGCACTGAGAAAGAAAGTATGCGACAGAAAGTACTGATGAGAGCTGACACCA 5917
 Qy 2644 eLysTyrSerLysAlaThrAspGlnGluLea-----TyrLe 2657
 Db 5918 GGTATACACACGAGATTTCCAGATTCGAGAGAGAGTACCCAGCGAGATTAAT-- 5975
 Qy 2657 uLysSerCysleuGlnAspLysGlnGlnGluLysuArgArgLysGlnLysLeuArgAr 2677
 Db 5976 ----ACACCTATGAGAGACAGTGAAGAAACCTCTCAGTCTTCACAAATAGATCTTGGCC 6031
 Qy 2677 gAlaGlnAlaAspAsnAspThrThr----- 2685
 Db 6032 ACTTCATCAGATCAGCAGACAGCAACTTCATCCAGATGCTCAAGGCAAGAGATGA 6091
 Qy 2686 ----ValCysValProLysAspTyrGlnLysAla----- 2695
 Db 6092 TGTCAATTGTAATTGACAGTATGATGAAGAGAGATGAGAAATGATGATGATGA 6151
 Qy 2695 ----- 2695
 Db 6152 AGATGACACAGGATGGAGATGAGGTTGAAGATAGTAATGAAGAACTGTAGTCCGA 6211
 Qy 2696 -----SerThrPheProValTh 2701
 Db 6212 TGGCAATGATGTTATGAAGCTGATGATGAGGCTGTGATGAGTACGATGACAGGATAC 6271
 Qy 2701 r-----CysGlyGlyLysGlyLysIleValGlnSerThrAlaMetLeuVa 2716
 Db 6272 AGAAACAGAGAAAGTATGAGTGGAGTGAAGCT----- 6305
 Qy 2716 lLeuGlnSerGluLysAlaAlaLeuGlnArgGlnLeuSerHisTyrLysLysTyrH 2736
 Db 6306 -----AATACAGAGCTGCTGATTCCAAAACAGTGTGAAGAAATACAGTGTGTC 6358
 Qy 2736 sHisLeuSerArgThrMetSerSerSerGluAspArgLysThrLysAlaLysSerAs 2756
 Db 6359 AGAATCTTCTTTTCTCAGAGAGTTCAGAGAAACAAACAGCATCATCAGCA---TCGA 6415
 Qy 2756 pAlaHisSerSerHisThrGlySerSerHisArgGlySerProHisLysThrGlnThTy 2776
 Db 6416 AAGACAGCGCCCTCGAGCACTCACTGACGACGAGACGCCACACAT----- 6461

```
OY 2776 rAIGHISGLyProValThrProGluArgSer--GluMetProSerLeuHisLeuGlySe 2795
Db 6462 -----CCACTCCCCCAAGACTGACATTCATGCCCCACCTCCAGAGATTGGAC 6511
OY 2795 rProLysLysSerGluSerSerThrLysArgValAlaSerProAsnArgSerGluIleTy 2815
Db 6512 ACCA-----GTTCAAGAAATTCAGATGACCCGAGCAGCTGTAGG 6553
OY 2815 rSerGluLeuValMetSerProGlyLysThrGlyMetHisLysHisLysLeuSer----- 2833
Db 6554 ACGGGCCTTCAGTTGACTCCAGGAATAGGTGGCATGCAACGATTTTGTGATGATGA 6613
OY 2834 -----ProSerLysValGlyLeuHisLysLysArgAlaLeuSerProAsnAr 2849
Db 6614 AGACAGAACAGTCCAGATTCACACTCTT-----GTGGTCCACATCG 6658
OY 2849 gSerGluMetProThrGlnHisValIleSerProGlyLysThrGlyLeuHisLys----- 2867
Db 6659 TACTGATGATTGCTGCAAGCAATTCATCCCGCAGGTTGCTGTGCTCCCTGATTCG 6718
OY 2868 -----AsnLeuThrGluSerThrLeuPheAspAsnLeuSerProCy 2882
Db 6719 GTTTGGCCACCTGAAGATATCCACAAACAAGT-----TCTAGTCACTGTGATCT 6769
OY 2882 sLysGlnGlnLysValGlnGluAsnLeuAsnSerProLysGlyLysLeuPhe----- 2899
Db 6770 TGGCCAGCTTGTCTCTCAAGAGAGGTTAGGAATGATGAAACACCCGTTCCTAGCTCA 6829
OY 2900 -----AspValLysSerLysSerMetProTyrcysProSerGln----- 2912
Db 6830 TGAAGAAGAGTCAGTGGCGGAGGTGCTCCACTACTCCACTACAAAGTAGACGCCCCAGT 6889
OY 2913 ----PhePheAspAsnSerLysLeuGlyAspPheSerGln 2924
Db 6890 GACTGTATTACTGAGAGCACCACTGTGATGCTTGGA 6929
```

Search completed: May 23, 2003, 19:43:25
Job time : 1712 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 23, 2003, 01:00:29 ; Search time 473 Seconds
(without alignments)
6241.164 Million cell updates/sec

Title: US-09-150-867-2

Perfect score: 9626
Sequence: 1 gattccgagtcgagtagy.....ttaaaaaaaaaacggaattc 9626

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 441362 seqs, 153338381 residues

Word size: 0

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: Issued_Patents_NA:*

1: /cgn2_6/prodata/2/lna/5A.COMB.seq:*
2: /cgn2_6/prodata/2/lna/5B.COMB.seq:*
3: /cgn2_6/prodata/2/lna/6A.COMB.seq:*
4: /cgn2_6/prodata/2/lna/6B.COMB.seq:*
5: /cgn2_6/prodata/2/lna/CTUS.COMB.seq:*
6: /cgn2_6/prodata/2/lna/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	0.2	10564	1 US-08-206-176-5	Sequence 5, Appl
2	19	0.2	423	3 US-08-906-769-134	Sequence 134, App
3	19	0.2	423	3 US-08-906-616-134	Sequence 134, App
4	19	0.2	423	3 US-08-639-075A-134	Sequence 134, App
5	19	0.2	423	4 US-09-012-431-134	Sequence 134, App
6	19	0.2	423	4 US-09-012-692-134	Sequence 134, App
7	19	0.2	423	4 US-08-906-613-134	Sequence 134, App
8	19	0.2	445	1 US-08-207-904-14	Sequence 14, Appl
9	19	0.2	658	4 US-09-328-111-665	Sequence 665, App
10	19	0.2	4080	1 US-08-375-300-1	Sequence 1, Appl
11	19	0.2	4080	1 US-09-177-431-1	Sequence 1, Appl
12	19	0.2	4080	5 US-09-155-16930-1	Sequence 1, Appl
13	19	0.2	21338	4 US-08-961-527-20	Sequence 20, Appl
14	19	0.2	42571	4 US-09-810-347-3	Sequence 3, Appl
15	18	0.2	31	4 US-08-859-998-170	Sequence 170, App
16	18	0.2	31	4 US-09-225-928-170	Sequence 170, App
17	18	0.2	147	4 US-09-269-833D-20	Sequence 20, Appl
18	18	0.2	315	2 US-08-743-200-1	Sequence 1, Appl
19	18	0.2	432	2 US-09-001-944-11	Sequence 11, Appl
20	18	0.2	432	4 US-09-240-004A-11	Sequence 11, Appl
21	18	0.2	435	2 US-09-001-944-9	Sequence 9, Appl
22	18	0.2	435	4 US-09-240-004A-9	Sequence 9, Appl
23	18	0.2	435	4 US-09-269-833D-30	Sequence 30, Appl
24	18	0.2	441	4 US-09-269-833D-28	Sequence 28, Appl
25	18	0.2	450	4 US-09-240-004A-13	Sequence 13, Appl
26	18	0.2	453	4 US-09-269-833D-29	Sequence 29, Appl
27	18	0.2	457	4 US-09-269-833D-26	Sequence 26, Appl

28	18	0.2	498	2 US-09-001-944-1	Sequence 1, Appl
29	18	0.2	498	2 US-09-001-944-3	Sequence 3, Appl
30	18	0.2	498	2 US-09-001-944-5	Sequence 5, Appl
31	18	0.2	498	2 US-09-001-944-7	Sequence 7, Appl
32	18	0.2	498	4 US-09-240-004A-1	Sequence 1, Appl
33	18	0.2	498	4 US-09-240-004A-3	Sequence 3, Appl
34	18	0.2	498	4 US-09-240-004A-5	Sequence 5, Appl
35	18	0.2	498	4 US-09-240-004A-7	Sequence 7, Appl
36	18	0.2	501	4 US-09-269-833D-3	Sequence 3, Appl
37	18	0.2	501	4 US-09-269-833D-27	Sequence 27, Appl
38	18	0.2	517	4 US-09-269-833D-13	Sequence 13, Appl
39	18	0.2	517	4 US-09-269-833D-23	Sequence 23, Appl
40	18	0.2	906	4 US-08-976-259-98	Sequence 98, Appl
41	18	0.2	955	4 US-09-641-638-9	Sequence 9, Appl
42	18	0.2	955	4 US-09-641-638-10	Sequence 10, Appl
43	18	0.2	955	4 US-09-641-638-11	Sequence 11, Appl
44	18	0.2	959	4 US-09-724-510-1	Sequence 1, Appl
45	18	0.2	959	4 US-09-723-216-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-206-176-5
; Sequence 5, Application US/08206176
; Patent No. 5639940
; GENERAL INFORMATION:
; APPLICANT: Garner, Ian
; APPLICANT: Dalrymple, Michael A
; APPLICANT: Prunkard, Donna E
; APPLICANT: Foster, Donald C
; TITLE OR INVENTION: Production of Fibrinogen in Transgenic
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 4225 Roosevelt Way, N.E.
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/206,176
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31-648
; REFERENCE/DOCKET NUMBER: 93-15
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-547-8080 ext 322
; TELEFAX: 206-548-2329
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10564 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: human fibrinogen gamma chain
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(1799..1876, 1973..2017, 2207..2390, 2510
; ..2603, 4211..4341, 4645..4778, 5758..5942, 7426
; ..7703, 9342..9571)
; LOCATION: ..7703, 9342..9571)
; US-08-206-176-5

Query Match 0.2%; Score 21; DB 1; Length 10564;
Best Local Similarity 100.0%; Pred. No. 8.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2975 TTCTCAAGCAGCTTGAGAAA 2995
|||||
Db 3246 TTCTCAAGCAGCTTGAGAAA 3266

RESULT 2

US-08-906-769-134
; Sequence 134, Application US/08906769
; Patent No. 6077687
; GENERAL INFORMATION:
; APPLICANT: Grieve, Robert B.
; APPLICANT: Rushlow, Keith E.
; APPLICANT: Wu Hunter, Shirley
; APPLICANT: Frank, Glenn R.
; APPLICANT: Stiegler, Gary
; APPLICANT: Gaines, Patrick J.
; TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
; TITLE OF INVENTION: MOLECULES AND USES THEREOF
; NUMBER OF SEQUENCES: 190
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/906,769
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/639,075
; FILING DATE: 24-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-25-C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 134:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 423 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..423
; US-08-906-769-134

Query Match 0.2%; Score 19; DB 3; Length 423;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1824 AACTACAGTCAAGAGAGA 1842
|||||
Db 222 AACTACAGTCAAGAGAGA 240

RESULT 3

US-08-906-616-134
; Sequence 134, Application US/08906616
; Patent No. 6121035
; GENERAL INFORMATION:
; APPLICANT: Grieve, Robert B.
; APPLICANT: Rushlow, Keith E.
; APPLICANT: Wu Hunter, Shirley
; APPLICANT: Frank, Glenn R.
; APPLICANT: Stiegler, Gary
; APPLICANT: Gaines, Patrick J.
; TITLE OF INVENTION: FLEA AMINOPEPTIDASE PROTEINS AND USES THEREOF
; NUMBER OF SEQUENCES: 190
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross P.C.
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/906,616
; FILING DATE: 05-AUG-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-25-C2-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 134:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 423 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..423
; US-08-906-616-134

Query Match 0.2%; Score 19; DB 3; Length 423;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1824 AACTACAGTCAAGAGAGA 1842
|||||
Db 222 AACTACAGTCAAGAGAGA 240

RESULT 4

US-08-639-075A-134
; Sequence 134, Application US/08639075A
; Patent No. 6150125
; GENERAL INFORMATION:
; APPLICANT: Grieve, Robert B.
; APPLICANT: Rushlow, Keith E.
; APPLICANT: Wu Hunter, Shirley
; APPLICANT: Frank, Glenn R.
; APPLICANT: Stiegler, Gary
; APPLICANT: Gaines, Patrick J.
; TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
; TITLE OF INVENTION: MOLECULES AND USES THEREOF
; NUMBER OF SEQUENCES: 190
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/639,075A
FILING DATE: 24-APR-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 134:
SEQUENCE CHARACTERISTICS:
LENGTH: 423 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..423
US-08-639-075A-134

Query Match 0.2%; Score 19; DB 3; Length 423;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1824 AACTACAGTCAAGAGGA 1842
DB 222 AACTACAGTCAAGAGGA 240

RESULT 5
US-09-012-431-134
Sequence 134, Application US/09012431
Patent No. 6180383
GENERAL INFORMATION:
APPLICANT: Grieve, Robert B.
Rushlow, Keith E.
Wu Hunter, Shirley
Frank, Glenn R.
Stiegler, Gary
Gaines, Patrick J.
Silver, Gary
TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
NUMBER OF SEQUENCES: 190
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/012,431
FILING DATE: 23-Jan-1998

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/639,075
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 134:
SEQUENCE CHARACTERISTICS:
LENGTH: 423 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..423
US-09-012-431-134

Query Match 0.2%; Score 19; DB 4; Length 423;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1824 AACTACAGTCAAGAGGA 1842
DB 222 AACTACAGTCAAGAGGA 240

RESULT 6
US-09-012-692-134
Sequence 134, Application US/09012692
Patent No. 6214579
GENERAL INFORMATION:
APPLICANT: Grieve, Robert B.
Rushlow, Keith E.
Wu Hunter, Shirley
Frank, Glenn R.
Stiegler, Gary
Gaines, Patrick J.
Silver, Gary
TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
NUMBER OF SEQUENCES: 190
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/012,692
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/639,075
FILING DATE: 24-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700

TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 134:
SEQUENCE CHARACTERISTICS:
LENGTH: 423 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..423
US-09-012-692-134

Query Match 0.2%; Score 19; DB 4; Length 423;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1824 AACTACAGTCAAGAGGA 1842
DB 222 AACTACAGTCAAGAGGA 240

RESULT 7
US-08-906-613-134
Sequence 134, Application US/08906613
Patent No. 6232096
GENERAL INFORMATION:
APPLICANT: Grieve, Robert B.
APPLICANT: Rushlow, Keith E.
APPLICANT: Wu Hunter, Shirley
APPLICANT: Frank, Glenn R.
APPLICANT: Stiegler, Gary
APPLICANT: Gaines, Patrick J.
TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
TITLE OF INVENTION: MOLECULES AND USES THEREOF
NUMBER OF SEQUENCES: 190
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/906,613
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/639,075
FILING DATE: 24-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 134:
SEQUENCE CHARACTERISTICS:
LENGTH: 423 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..423

US-08-906-613-134

Query Match 0.2%; Score 19; DB 4; Length 423;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1824 AACTACAGTCAAGAGGA 1842
DB 222 AACTACAGTCAAGAGGA 240

RESULT 8
US-08-207-904-14
Sequence 14, Application US/08207904
Patent No. 5477002
GENERAL INFORMATION:
APPLICANT: Tuttle, AnnMarie
APPLICANT: Crossland, Lyle D.
TITLE OF INVENTION: Another-Specific CDNA Sequences, Genomic
TITLE OF INVENTION: DNA Sequences and Recombinant DNA Sequences
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESSES:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: New York
COUNTRY: USA
ZIP: 10532

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/207,904
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/908,242
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Lazar, Steven R.
REGISTRATION NUMBER: 32,618
REFERENCE/DOCKET NUMBER: CGC 1624
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8615
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 445 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Nicotiana tabacum
INDIVIDUAL ISOLATE: Ant68
FEATURE:
NAME/KEY: CDS
LOCATION: 2..445
US-08-207-904-14

Query Match 0.2%; Score 19; DB 1; Length 445;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4924 CAATGTCGAATCACTCA 4942
DB 301 CAATGTCGAATCACTCA 319

RESULT 9
US-09-328-111-665
Sequence 665, Application US/09328111
Patent No. 6262333
GENERAL INFORMATION:
APPLICANT: Endege, Wilson O.
APPLICANT: Steinmann, Kathleen E.
APPLICANT: Astle, Jon H.
APPLICANT: Burgess, Christopher C.
APPLICANT: Bushnell, Steven E.
APPLICANT: Carroll III, Eddie
APPLICANT: Catino, Theodore J.
APPLICANT: Derli, Adnan
APPLICANT: Ford, Donna M.
APPLICANT: Lewis, Marcia E.
APPLICANT: Monahan, John E.
APPLICANT: Schlegel, Robert
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
FILE REFERENCE: CCD-257 (US)
CURRENT APPLICATION NUMBER: US/09/328,111
EARLIER FILING DATE: 1998-06-08
NUMBER OF SEQ ID NOS: 850
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 665
LENGTH: 658
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc-feature
LOCATION: (1)...(658)
OTHER INFORMATION: n - A,T,C or G
US-09-328-111-665

Query Match 0.2%; Score 19; DB 4; Length 658;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3813 TGCAGCTGGAGATCTGCA 3831
DB 24 TGCAGCTGGAGATCTGCA 42

RESULT 10
US-08-375-300-1
Sequence 1, Application US/08375300
Patent No. 5679566
GENERAL INFORMATION:
APPLICANT: Feng, He
APPLICANT: Jacobson, Allan S.
TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE PRODUCTION IN
THE ABSENCE OF NONSENSE-MEDIATED mRNA DECAY FUNCTION
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street Suite 3100
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/375,300
FILING DATE: 20-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Fasse, J. P.
REGISTRATION NUMBER: 32,983

REFERENCE/DOCKET NUMBER: 04020/046001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)542-5070
TELEFAX: (617)542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4080 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-375-300-1

Query Match 0.2%; Score 19; DB 1; Length 4080;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4166 GAAGAGATGAGCTTCAAA 4184
DB 862 GAAGAGATGAGCTTCAAA 880

RESULT 11
US-09-177-431-1
Sequence 1, Application US/09177431
Patent No. 6071700
GENERAL INFORMATION:
APPLICANT: He, Feng
APPLICANT: Jacobson, Allan S.
TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE PRODUCTION IN THE
ABSENCE OF NONSENSE-MEDIATED mRNA DECAY FUNCTION
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/177,431
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/955,472
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Fasse, J. Peter
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 07917/050001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-9806
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4080 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-177-431-1

Query Match 0.2%; Score 19; DB 3; Length 4080;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4166 GAAGAGATGAGCTTCAAA 4184

Db 862 GAAAGAGATGAGCTTCAA 880

RESULT 12
PCT-US95-16930-1
Sequence 1, Application PC/TUS9516930

GENERAL INFORMATION:
APPLICANT: UNIVERSITY OF MASSACHUSETTS MEDICAL SCHOOL
TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE
TITLE OF INVENTION: PRODUCTION IN THE ABSENCE OF
TITLE OF INVENTION: NONSENSE-MEDIATED MRNA DECAY
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street Suite 3100
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/16930
FILING DATE: 27-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/375,300
FILING DATE: 20-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Fasse, J. Peter
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 04020/046MO1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)542-5070
TELEFAX: (617)542-8906
TELEX: 200154

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4080 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
PCT-US95-16930-1

Query Match 0.2%; Score 19; DB 5; Length 4080;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4166 GAAAGAGATGAGCTTCAA 4184
Db 862 GAAAGAGATGAGCTTCAA 880

RESULT 13
US-08-961-527-20
Sequence 20, Application US/08961527
Patent No. 6420135
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
NUMBER OF SEQUENCES: 391
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 Inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,527
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8512
TELEFAX: (301) 309-8504
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 21338 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-527-20

Query Match 0.2%; Score 19; DB 4; Length 21338;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3446 AAAAAATCTTTTTCAG 3464
Db 1768 AAAAAATCTTTTTCAG 1786

RESULT 14
US-09-810-347-3
Sequence 3, Application US/09810347
Patent No. 6461847
GENERAL INFORMATION:
APPLICANT: Ye, Jane et al.
TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
FILE REFERENCE: C1001169
CURRENT APPLICATION NUMBER: US/09/810,347
CURRENT FILING DATE: 2001-03-19
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 42571
TYPE: DNA
ORGANISM: Human
US-09-810-347-3

Query Match 0.2%; Score 19; DB 4; Length 42571;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3109 GATTAATGTTATTAAGCA 3127
Db 39653 GATTAATGTTATTAAGCA 39671

RESULT 15
US-08-859-998-170
Sequence 170, Application US/08859998
Patent No. 5994076
GENERAL INFORMATION:
APPLICANT: Chenchik, Alex
APPLICANT: Johadze, George
APPLICANT: Biblashvili, Robert

```

: TITLE OF INVENTION: METHOD OF ASSAYING DIFFERENTIAL
: TITLE OF INVENTION: EXPRESSION
: NUMBER OF SEQUENCES: 1375
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson, P.C.
: STREET: 2200 Sand Hill Road, Suite 100
: CITY: Menlo Park
: STATE: CA
: COUNTRY: US
: ZIP: 94025
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: Windows95
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/859,998
: FILING DATE: 21-MAY-1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Field, Bret E.
: REGISTRATION NUMBER: 37,620
: REFERENCE/DOCKET NUMBER: 09096/002001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-322-5070
: TELEFAX: 415-854-0875
: INFORMATION FOR SEQ ID NO: 170:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 31 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA
: FEATURE:
: OTHER INFORMATION: oligonucleotide primer
:
: US-08-859-998-170

```

Query Match 0.28; Score 18; DB 2; Length 31;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5338 ATGCTGCTAATGAGAA 5355
 |||
 DB 7 ATGCTGCTAATGAGAA 24

Search completed: May 23, 2003, 08:05:39
 Job time : 910 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 23, 2003, 07:32:30 ; Search time 1097 Seconds
(Without alignments)
11386.846 Million cell updates/sec

Title: US-09-150-867-2

Perfect score: 9626

Sequence: 1 gaattccgagtcgtagg.....ttaaaaaaaaaacggaattc 9626

Scoring table:

OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 828747 seqs, 660231138 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1657494

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Published_Applications_NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCr_NEM_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEM_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US07_NEM_PUB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US08_NEM_PUB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEM_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEM_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEM_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	24	0.2	1100	10 US-09-263-959-276	Sequence 276, App
C 2	24	0.2	1277	10 US-09-263-959-269	Sequence 269, App
C 3	24	0.2	640681	10 US-09-790-988-1	Sequence 1, Appl
C 4	24	0.2	684973	10 US-09-263-959-1	Sequence 1, Appl
C 5	23	0.2	7684	10 US-09-070-927A-16	Sequence 16, Appl
C 6	20	0.2	1215	9 US-09-925-299-79	Sequence 79, Appl
C 7	20	0.2	1215	10 US-09-925-299-79	Sequence 79, Appl
C 8	20	0.2	4510	12 US-10-005-983-1	Sequence 1, Appl
C 9	20	0.2	7992	9 US-09-893-519A-140	Sequence 140, App
C 10	20	0.2	8493	9 US-10-071-766-51	Sequence 51, Appl
C 11	20	0.2	180557	12 US-10-003-806-6	Sequence 6, Appl
C 12	20	0.2	180557	12 US-10-003-806-6	Sequence 6, Appl
C 13	20	0.2	402850	9 US-09-844-653-5	Sequence 5, Appl
C 14	19	0.2	70	10 US-09-815-242-371	Sequence 371, App
C 15	19	0.2	231	10 US-09-864-761-31477	Sequence 31477, A
C 16	19	0.2	247	9 US-09-918-995-18268	Sequence 18268, A
C 17	19	0.2	344	9 US-09-918-995-29936	Sequence 29936, A
C 18	19	0.2	348	10 US-09-864-761-14056	Sequence 14056, A
C 19	19	0.2	355	10 US-09-864-761-28258	Sequence 28258, A

20	19	0.2	358	10 US-09-960-352-13853	Sequence 13853, A
C 21	19	0.2	361	10 US-09-864-761-22536	Sequence 22536, A
C 22	19	0.2	364	10 US-09-783-590-5629	Sequence 5629, App
C 23	19	0.2	388	9 US-09-918-995-3711	Sequence 3711, App
C 24	19	0.2	430	9 US-09-918-995-34048	Sequence 34048, A
C 25	19	0.2	432	9 US-09-918-995-34983	Sequence 34983, A
C 26	19	0.2	440	10 US-09-960-352-9139	Sequence 9139, App
C 27	19	0.2	443	10 US-09-764-877-900	Sequence 900, App
C 28	19	0.2	455	10 US-09-864-761-11679	Sequence 11679, A
C 29	19	0.2	457	10 US-09-864-761-5775	Sequence 5775, App
C 30	19	0.2	577	10 US-09-864-761-14948	Sequence 14948, A
C 31	19	0.2	591	10 US-09-864-761-7654	Sequence 7654, App
C 32	19	0.2	639	9 US-10-066-543-666	Sequence 666, App
C 33	19	0.2	658	10 US-09-879-536-665	Sequence 665, App
C 34	19	0.2	740	10 US-09-910-843-690	Sequence 690, App
C 35	19	0.2	771	12 US-10-001-843-64	Sequence 64, Appl
C 36	19	0.2	1100	10 US-09-263-959-275	Sequence 275, App
C 37	19	0.2	1160	9 US-09-764-891-10183	Sequence 10183, A
C 38	19	0.2	1203	10 US-09-263-959-239	Sequence 239, App
C 39	19	0.2	1381	9 US-10-023-282-82	Sequence 82, Appl
C 40	19	0.2	1439	10 US-09-882-830A-51	Sequence 51, Appl
C 41	19	0.2	1728	9 US-09-938-842A-2550	Sequence 2550, App
C 42	19	0.2	1838	9 US-09-938-842A-4073	Sequence 4073, App
C 43	19	0.2	1880	10 US-09-731-872-78	Sequence 78, Appl
C 44	19	0.2	2000	9 US-09-938-842A-4182	Sequence 4182, App
C 45	19	0.2	2109	10 US-09-070-927A-488	Sequence 488, App

ALIGNMENTS

RESULT 1

US-09-263-959-276/C
Sequence 276, Application US/09263959
Patent No. US20020150891A1

GENERAL INFORMATION:

APPLICANT: Hoood, Leroy E.

APPLICANT: Rowen, Lee

APPLICANT: Koop, Ben F.

TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC COMPOSITIONS AND METHODS WHICH U

NUMBER OF SEQUENCES: 1279

CORRESPONDENCE ADDRESS:

ADDRESSEE: Seed and Berry LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: US

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/263, 959

FILING DATE: 05-MAR-1999

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: McMaisters, David D.

REGISTRATION NUMBER: 33,963

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 276:

SEQUENCE CHARACTERISTICS:

LENGTH: 1100 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-09-263-959-276

Query Match 0.2%; Score 24; DB 10; Length 1100;

Best Local Similarity 100.0%; Pred. No. 0.56;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5855 GCAGAGAGCAAAAGCTGATGCT 5878
|||||
Db 689 GCAGAGAGCAAAAGCTGATGCT 666

RESULT 2

US-09-263-959-269/C
; Sequence 269, Application US/09263959
; Patent No. US20020150891A1
; GENERAL INFORMATION:
; APPLICANT: Hood, Leroy E.
; APPLICANT: Rowen, Lee
; APPLICANT: Koop, Ben F.
; TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC COMPOSITIONS AND METHODS WHICH UTILIZE
; NUMBER OF SEQUENCES: 1279
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/263,959
; FILING DATE: 05-MAR-1999
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mcmasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 920010.426C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 269:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1277 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-263-959-269

Query Match 0.2%; Score 24; DB 10; Length 1277;

Best Local Similarity 100.0%; Pred. No. 0.57;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5855 GCAGAGAGCAAAAGCTGATGCT 5878
|||||
Db 758 GCAGAGAGCAAAAGCTGATGCT 735

RESULT 3

US-09-790-988-1
; Sequence 1, Application US/09790988
; Patent No. US20020127687A1
; GENERAL INFORMATION:
; APPLICANT: SHIGENOBU, SHUJI
; APPLICANT: WATANABE, HIDEMI
; APPLICANT: HATORI, MASAHIRA
; APPLICANT: SAKAKI, YOSHIYUKI
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
; FILE REFERENCE: 081356/0159
; CURRENT APPLICATION NUMBER: US/09/790,988
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: JP2000-107160
; PRIOR FILING DATE: 2000-04-07

; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1

; LENGTH: 640681

; TYPE: DNA

; ORGANISM: Buchnera sp.

US-09-790-988-1

Query Match 0.2%; Score 24; DB 10; Length 640681;
Best Local Similarity 100.0%; Pred. No. 0.93;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3957 AAAACATTGAGCAGTCATTTGAA 3980
|||||
Db 505797 AAAACATTGAGCAGTCATTTGAA 505820

RESULT 4

US-09-263-959-1/C
; Sequence 1, Application US/09263959
; Patent No. US20020150891A1
; GENERAL INFORMATION:
; APPLICANT: Hood, Leroy E.
; APPLICANT: Rowen, Lee
; APPLICANT: Koop, Ben F.
; TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC COMPOSITIONS AND METHODS WHICH UTILIZE
; NUMBER OF SEQUENCES: 1279
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/263,959
; FILING DATE: 05-MAR-1999
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mcmasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 920010.426C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 684973 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-263-959-1

Query Match 0.2%; Score 24; DB 10; Length 684973;

Best Local Similarity 100.0%; Pred. No. 0.94;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5855 GCAGAGAGCAAAAGCTGATGCT 5878
|||||
Db 163714 GCAGAGAGCAAAAGCTGATGCT 163691

RESULT 5

US-09-070-927A-16
; Sequence 16, Application US/09070927A
; Patent No. US20020120116A1
; GENERAL INFORMATION:
; APPLICANT: Charles A. Kunsch

Patrick J. Dillon
Steven Barash
TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 982
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44M storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/070,927A
FILING DATE: 04-May-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/046,655
FILING DATE: 1997-05-16
APPLICATION NUMBER: 60/044,031
FILING DATE: 1997-05-06
APPLICATION NUMBER: 60/066,009
FILING DATE: 1997-11-14
ATTORNEY/AGENT INFORMATION:
NAME: Kenley K. Hoover
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: PB369
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 7684 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-09-070-927A-16
Query Match 0.2%; Score 23; DB 10; Length 7684;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3973 AATTGAACACAGATGAATTAA 3995
|||||
DB 4287 AATTGAACACAGATGAATTAA 4309
RESULT 6
US-09-925-299-79
Sequence 79, Application US/09925299
Publication No. US20030040617A9
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA102
CURRENT APPLICATION NUMBER: US/09/925,299
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05883
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1556
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 79
LENGTH: 1215
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-299-79

Query Match 0.2%; Score 20; DB 9; Length 1215;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 9599 ATTTAAATTAAAAAAA 9618
|||||
DB 1178 ATTTAAATTAAAAAAA 1197
RESULT 7
US-09-925-299-79
Sequence 79, Application US/09925299
Patent No. US20020055627A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA102
CURRENT APPLICATION NUMBER: US/09/925,299
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05883
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1556
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 79
LENGTH: 1215
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-299-79
Query Match 0.2%; Score 20; DB 10; Length 1215;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 9599 ATTTAAATTAAAAAAA 9618
|||||
DB 1178 ATTTAAATTAAAAAAA 1197
RESULT 8
US-10-005-983-1
Sequence 1, Application US/10005983
Patent No. US20020116730A1
GENERAL INFORMATION:
APPLICANT: Allen, Keith D.
TITLE OF INVENTION: TRANSGENIC MICE CONTAINING PERK PROTEIN
FILE REFERENCE: R-517
CURRENT APPLICATION NUMBER: US/10/005,983
CURRENT FILING DATE: 2001-11-07
PRIOR APPLICATION NUMBER: US 60/246,676
PRIOR FILING DATE: 2000-11-07
PRIOR APPLICATION NUMBER: US 60/311,018
PRIOR FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/324,765
PRIOR FILING DATE: 2001-09-24
PRIOR APPLICATION NUMBER: US 60/326,148
PRIOR FILING DATE: 2001-09-28
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 4510
TYPE: DNA
ORGANISM: Mus musculus
US-10-005-983-1
Query Match 0.2%; Score 20; DB 12; Length 4510;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 9599 ATTTAAATTAAAAAAA 9618

Db 4476 ATTATTAAGAAAAA 4495

RESULT 9
US-09-893-519A-140

; Sequence 140, Application US/09893519A
; Publication No. US20030027243A1
; GENERAL INFORMATION:
; APPLICANT: ANADYS PHARMACEUTICALS, INC.
; APPLICANT: THOMPSON, Craig
; APPLICANT: MOORE, Jeffrey
; APPLICANT: BURMAN, Ed T.
; APPLICANT: BRADLEY, John
; APPLICANT: DESTIVA, Thamar
; APPLICANT: HARRIS, Sandra
; APPLICANT: KOMARNITSKY, Svetlana
; APPLICANT: MENDILLO, Marc
; APPLICANT: MCCOY, Daniel
; APPLICANT: SANDERSON, Karen
; APPLICANT: HAO, Tariq
; APPLICANT: ZHU, Shuhao
; APPLICANT: LONG, Fan
; APPLICANT: DAVIDOV, Eugene
; TITLE OF INVENTION: ANTIFUNGAL COMPOUNDS AND METHODS OF USE
; FILE REFERENCE: 0342/1G348-US2
; CURRENT FILING DATE: 2001-06-28
; PRIOR FILING DATE: 2000-06-29
; PRIOR FILING DATE: 2000-06-29
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 146
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 140
; LENGTH: 7992
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: Human Genbank/Z15005.1
; DATABASE ENTRY DATE: 1993-01-10
; RELEVANT RESIDUES: (1)..(7992)
US-09-893-519A-140

Query Match 0.2%; Score 20; DB 9; Length 7992;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8012 CCATTAAGAAATTTGA 8031
Db 7240 CCATTAAGAAATTTGA 7259

RESULT 10
US-10-071-766-51
; Sequence 51, Application US/10071766
; Publication No. US20020192678A1
; GENERAL INFORMATION:
; APPLICANT: Huet-Mei Chen
; TITLE OF INVENTION: GENES EXPRESSED IN SENESENCE
; FILE REFERENCE: PA-0043 US
; CURRENT APPLICATION NUMBER: US/10/071,766
; CURRENT FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: PERL Program
; SEQ ID NO 51
; LENGTH: 8493
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020192678A1 441283.3

US-10-071-766-51

Query Match 0.2%; Score 20; DB 9; Length 8493;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8012 CCATTAAGAAATTTGA 8031
Db 7330 CCATTAAGAAATTTGA 7349

RESULT 11

US-10-003-806-6/c
; Sequence 6, Application US/10003806
; Patent No. US20020119929A1
; GENERAL INFORMATION:
; APPLICANT: Bishop, Colin E.
; APPLICANT: Agoulrik, Alexander I.
; TITLE OF INVENTION: CAN 1 AND ITS ROLE IN MAMMALIAN INFERTILITY
; FILE REFERENCE: P02066US1/10024824
; CURRENT APPLICATION NUMBER: US/10/003,806
; CURRENT FILING DATE: 2001-11-02
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 180557
; TYPE: DNA
; ORGANISM: Human
US-10-003-806-6

Query Match 0.2%; Score 20; DB 12; Length 180557;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5449 AAAAGATCTGAACAAAAG 5468
Db 91760 AAAAGATCTGAACAAAAG 91741

RESULT 12

US-10-003-806-9/c
; Sequence 9, Application US/10003806
; Patent No. US20020119929A1
; GENERAL INFORMATION:
; APPLICANT: Bishop, Colin E.
; APPLICANT: Agoulrik, Alexander I.
; TITLE OF INVENTION: CAN 1 AND ITS ROLE IN MAMMALIAN INFERTILITY
; FILE REFERENCE: P02066US1/10024824
; CURRENT APPLICATION NUMBER: US/10/003,806
; CURRENT FILING DATE: 2001-11-02
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 180557
; TYPE: DNA
; ORGANISM: Human
US-10-003-806-9

Query Match 0.2%; Score 20; DB 12; Length 180557;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5449 AAAAGATCTGAACAAAAG 5468
Db 91760 AAAAGATCTGAACAAAAG 91741

```
RESULT 13
US-09-844-653-5
; Sequence 5, Application US/09844653
; Publication No. US20030054347A1
; GENERAL INFORMATION:
; APPLICANT: Richards, Julia
; APPLICANT: Rozsa, Frank
; TITLE OF INVENTION: Detecting and Treating Eye Disease
; FILE REFERENCE: US-06105
; CURRENT APPLICATION NUMBER: US/09/844,653
; CURRENT FILING DATE: 2001-04-27
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: Patent version 3.0
; SEQ ID NO 5
; LENGTH: 402850
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (46565)..(46565)
; OTHER INFORMATION: The n at this position can be a, c, t, or g.
; NAME/KEY: misc-feature
; LOCATION: (117443)..(118256)
; OTHER INFORMATION: The n at these positions can be a, c, t, or g.
; NAME/KEY: misc-feature
; LOCATION: (118272)..(118736)
; OTHER INFORMATION: The n at these positions can be a, c, t, or g.
; NAME/KEY: misc-feature
; LOCATION: (120172)..(120640)
; OTHER INFORMATION: The n at these positions can be a, c, t, or g.
; NAME/KEY: misc-feature
; LOCATION: (122654)..(122654)
; OTHER INFORMATION: The n at this position can be a, c, t, or g.
; NAME/KEY: misc-feature
; LOCATION: (203527)..(203963)
; OTHER INFORMATION: The n at these positions can be a, c, t, or g.
; NAME/KEY: misc-feature
; LOCATION: (215340)..(215340)
; OTHER INFORMATION: The n at this position can be a, c, t, or g.
; NAME/KEY: misc-feature
; LOCATION: (222415)..(222416)
; OTHER INFORMATION: The n at these positions can be a, c, t, or g.
; OTHER INFORMATION: The n at these positions can be a, c, t, or g.
US-09-844-653-5

Query Match          0.2%; Score 20; DB 9; Length 402850;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1834 AAAGAGGAGAGAGAGAGAGAG 1853
DB 300865 AAAGAGGAGAGAGAGAGAGG 300884

RESULT 14
US-09-815-242-371
; Sequence 371, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlson, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: E1179A.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
```

```
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: fastSeq for Windows Version 4.0
; SEQ ID NO 371
; LENGTH: 70
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (1)...(70)
; OTHER INFORMATION: n = A,T,C or G
US-09-815-242-371

Query Match          0.2%; Score 19; DB 10; Length 70;
Best Local Similarity 100.0%; Pred. No. 1,56+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1819 GCACACACTACAGTCAAG 1837
DB 51 GCACACACTACAGTCAAG 69

RESULT 15
US-09-864-761-31477
; Sequence 31477, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
```

Sun May 25 14:36:33 2003

us-09-150-867-2.01i.rmpb

```
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 31477
; LENGTH: 231
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AB016897.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
; OTHER INFORMATION: NT HIT: AB011399.1, EVALU 1.00e-127
; OTHER INFORMATION: EST HUMAN HIT: BF528927.1, EVALU 2.10e-01
; OTHER INFORMATION: SWISSPROT HIT: Q59146, EVALU 3.60e-01
; US-09-864-761-31477

Query Match          0.2%; Score 19; DB 10; Length 231;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1876 CAAGATGCGAGAACTGGA 1894
   |||
Db 176 CAAGATGCGAGAACTGGA 194

Search completed: May 23, 2003, 18:39:49
Job time : 14864 secs
```

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 22, 2003, 22:45:04 ; Search time 471 Seconds
(without alignments)
6267.666 Million cell updates/sec

Title: US-09-150-867-2

Perfect score: 9626
Sequence: 1 gaattccggagtcgagatag.....ttaaaaaaaaaacgaattc 9626

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA: *
1: /cgn2_6/ptodata/2/1na/5A_COMB.seq: *
2: /cgn2_6/ptodata/2/1na/5B_COMB.seq: *
3: /cgn2_6/ptodata/2/1na/6A_COMB.seq: *
4: /cgn2_6/ptodata/2/1na/6B_COMB.seq: *
5: /cgn2_6/ptodata/2/1na/PCrus_COMB.seq: *
6: /cgn2_6/ptodata/2/1na/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	147.4	1.5	3837	US-09-724-517-1	Sequence 1, Appl1
2	147.4	1.5	3837	US-09-641-807A-1	Sequence 1, Appl1
3	147.4	1.5	3837	US-09-723-096-1	Sequence 1, Appl1
4	140	1.5	1023	US-09-724-517-3	Sequence 3, Appl1
5	140	1.5	1023	US-09-641-807A-3	Sequence 3, Appl1
6	140	1.5	1023	US-09-723-096-3	Sequence 3, Appl1
7	135.2	1.4	1421	US-09-592-054-5	Sequence 5, Appl1
8	135.2	1.4	1569	US-09-592-054-3	Sequence 3, Appl1
9	127.2	1.3	4127	US-09-592-054-7	Sequence 7, Appl1
10	116	1.2	4308	US-09-592-054-1	Sequence 1, Appl1
11	98	1.0	1011	US-09-641-806-3	Sequence 3, Appl1
12	98	1.0	1026	US-09-641-806-1	Sequence 1, Appl1
13	95.8	1.0	1542	US-09-724-519-5	Sequence 5, Appl1
14	95.8	1.0	1542	US-09-592-037-5	Sequence 5, Appl1
15	95.8	1.0	1728	US-09-724-519-7	Sequence 7, Appl1
16	95.8	1.0	1728	US-09-592-037-7	Sequence 7, Appl1
17	95.8	1.0	3741	US-09-541-782-9	Sequence 9, Appl1
18	95.8	1.0	3741	US-09-723-820-9	Sequence 9, Appl1
19	90	0.9	29793	US-09-302-812-38	Sequence 38, Appl1
20	90	0.9	29793	US-09-511-477-38	Sequence 38, Appl1
21	90	0.9	29793	US-09-511-507-38	Sequence 38, Appl1
22	88	0.9	5093	US-08-468-036-23	Sequence 23, Appl1
23	88	0.9	5093	US-08-376-843-23	Sequence 23, Appl1
24	86.2	0.9	1107	US-09-724-519-9	Sequence 9, Appl1
25	86.2	0.9	1107	US-09-592-037-9	Sequence 9, Appl1
26	85.2	0.9	7218	US-08-232-463-14	Sequence 14, Appl1
27	82.8	0.9	1077	US-09-722-139-3	Sequence 3, Appl1

28	82.8	0.9	1077	US-09-721-832-3	Sequence 3, Appl1
29	82.8	0.9	1077	US-09-721-689-3	Sequence 3, Appl1
30	82.8	0.9	4176	US-09-722-139-1	Sequence 1, Appl1
31	82.8	0.9	4176	US-09-721-832-1	Sequence 1, Appl1
32	82.8	0.9	4176	US-09-721-689-1	Sequence 1, Appl1
33	82	0.9	1230	US-09-572-191-5	Sequence 5, Appl1
34	82	0.9	1230	US-09-723-262-5	Sequence 5, Appl1
35	82	0.9	1230	US-09-723-219-5	Sequence 5, Appl1
36	82	0.9	4757	US-09-572-191-1	Sequence 1, Appl1
37	82	0.9	4757	US-09-723-262-1	Sequence 1, Appl1
38	82	0.9	4757	US-09-723-219-1	Sequence 1, Appl1
39	79.8	0.8	2389	US-08-480-553-18	Sequence 18, Appl1
40	79.8	0.8	2389	US-08-486-382-4	Sequence 4, Appl1
41	79.8	0.8	2389	US-09-235-546-4	Sequence 4, Appl1
42	79.8	0.8	2389	US-08-923-208-18	Sequence 18, Appl1
43	79.8	0.8	2389	US-09-568-315-18	Sequence 18, Appl1
44	79.2	0.8	1149	US-09-724-519-3	Sequence 3, Appl1
45	79.2	0.8	1149	US-09-592-037-3	Sequence 3, Appl1

ALIGNMENTS

```
RESULT 1
US-09-724-517-1
; Sequence 1, Application US/09724517
; Patent No. 6379941
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6379941el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1031
; CURRENT APPLICATION NUMBER: US/09/724,517
; CURRENT FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: US/09/641,807
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3837
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (126)...(1337)
; OTHER INFORMATION: n = a, t, c, or g
US-09-724-517-1
Query Match 1.5%; Score 147.4; DB 4; Length 3837;
Best Local Similarity 51.3%; Pred No. 1.6e-26;
Matches 467; Conservative 0; Mismatches 411; Indels 33; Gaps 4;
275 TTCAAATTCGATCGGTATTAATCTCAGATCAACAGTCAATTTACGAGAAATA 334
136 TTCATCTGATTTGTTGTTTGGCAAAATTCACATCAAGATGAAGTATTAACATGT 195
335 GCAGTACATCTATATACGATCAGCTTTCGAGGATATATGACATATTTGATACGGA 394
196 ATTAAGCCCCCTGTGTGTCACACTGAGGGCTTAATGCAACGTTTCCCTATGGA 255
395 CAGCATCTTCAGGCAAGCATACATGATGAGGAACCAAAATTCATT----- 444
256 CAAATGATCTGCGAAGCATACACATTTGAGGGGCCATATTCCTCAGTTGGAG 315
445 -----GGGATATATACCCCAAGCATACAGAAAGTTTAAATTTATTCAGAGATA 496
316 GGCAGAAAGGATATCTTCCTCGAGCTATTCAGAAATTTTCAAGATCTGTGAAT 375
497 CCGAAGAGAGATTTCTTTAAGAGTTTATATGAGATTTTACATGAATAACTGTGAAA 556
376 CCTAGCATGACTTATATGATAAAGATCTTATATGAAGTGTACAGAGAACCTTAAGA 435
```

```

QY 557 GACCTACTGTGTATGACAGAGAAAGAGCCCTGGAAATTCGAGATTTTAATAGA 616
Db 436 GATCTTTAGATTTGGAGACATTCATGAAAGATCTTCATCCGAGAAATGAAAGAA 495
QY 617 AACGTGATGTGTGACCTGACAGAACTTTGTAATGTTCTGGAACATGTAATACAG 676
Db 496 AACACAGTGTGTTGGGGCCAGAGAAATGCCATGTGAGAGTGAAGTGAATGATAGT 555
QY 677 TGGATCAAAAAGGTGAAAAAAGACATTTATGAGAGATTAATGATCATAGT 736
Db 556 CTTTGGAGATGGGGAATGACGACATACAGTACACTCAATGTAATGAGACATCC 615
QY 737 AGTCGTTACATACATATTTAGATGATGTTGTAAGCCGAGACAGAAAT-----GAT 790
Db 616 AGCAGATCATGCAATTTTACATCAGCATTTGTCAAGTTCAATAAATATGAGAGCA 675
QY 791 CCCCAATTCAGAGAACTGTGATGAGAGCTGTCAATGGATCTCATGTAATTTGATAGT 850
Db 676 GCTGAAAGATGATATGATGATTTCCCTCGGCATATTTGTCAAAAGTTCCACTTTGTGAT 735
QY 851 CTTGCTGCAAGTGAAGAGCAAGCCAACTGAGCTGAGAGTGTGAGACTTAAGAGAGC 910
Db 736 TTGGCAGAGATCAGAAAGATACCAAAACGAGGAAATAGTGTGACGTTCAAGAAATCC 795
QY 911 TGCACATCAACGCGACCTGTTTATCCCTGACAGGTTATTAAGAACTTAGGAGC-- 968
Db 796 ATTCAATCAATAGTATGATGCTGCTTTAGCAATGTAATTAAGCGCTCTGGAGACCA 855
QY 969 -GCCAGCTGTGATTTATTAATACAGAGCAAACTCAGCAATTTCTCAAAAT 1027
Db 856 CGCAGGAAGATTCATATTCATATTAAGGATGCTAAATTAACCGGCTTCTGAAGAT 915
QY 1028 TCATTGGAGAGAAATGCTAAACGTTATTAATTTGACAAATTAAGCCAGTTTC-----T 1081
Db 916 TCCTGGAGAGAGATGCTAAGATGTCATGATCAGATGTGACAGCCCTCCTCGAAT 975
QY 1082 TTTGATGAGACTCTAAATACACTTTCAGTTGCCAGTACTGCAAAACATGTGAGAAATACT 1141
Db 976 TTTGATGAGTCTTAATTTCTCTCAATATATGCCAAGAGACGAGAACTTAGAAACAA 1035
QY 1142 CCCCATGTTAA 1152
Db 1036 CCCACTGTAA 1046

```

```

RESULT 2
US-09-641-807A-1
; Sequence 1, Application us/09641807A
; Patent No. 6440731
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6440731el motor proteins and methods for
; FILE OF INVENTION: their use
; CURRENT APPLICATION NUMBER: us/09/641,807A
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3837
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc_feature
; LOCATION: (1226)....(1336)
; OTHER INFORMATION: n = a, t, c, or g
US-09-641-807A-1

```

```

Query Match 1.5%; Score 147.4; DB 4; Length 3837;
Best Local Similarity 51.3%; Pred. No. 1.6e-26;
Matches 467; Conservative 0; Mismatches 411; Indels 33; Gaps 4;

```

```

QY 275 TTCAATTCGATCCGTATTTAATTTCTCAGCAATCAACAGTCAATTTACAGAAATA 334
Db 136 TTCACTTTGATTTTGTGTTTGGCAAAATTCACCTCAAGATGAAATTTAACAATGT 195
QY 335 GCACTACCTATCATACATCAGTCTTTGCAAGGATATATGACACAAATTTGCTACGCA 394
Db 196 ATAAAGCCCCAGTGTGTCTACTCATCTGAGGCTATATGCACTGTTTGTGCTATGGA 255
QY 395 CAGACATCTTCAGCAAGACGTACACATATGATGGGAACCAAAATTCAT----- 444
Db 256 CAACATGATCTGGGAAGACATACACATTTGAGAGGGCCATATTCCTTCACTGTGAG 315
QY 445 -----GGGCATTAATACCCCAAGCCATACAGAAATTTTAAATATTCAGAGATA 496
Db 316 GGCCAAAAGGATATCATCTCCGAGATTTCAAGAAATTTCAACATCTGTGACAT 375
QY 497 CCGAAGAGAGTTCTTCTTAAGATTTCTTATATGAGATTTCAATGAACTGTGAAA 556
Db 376 CCTAGCATTTGACTTTAATGTAAGATCTTATATAGAGTCAAGAGAACCTAAGA 435
QY 557 GACCTACTGTGTATGACAGAGAAAGAGCCCTGGAAATTCGCGAGATTTTAATAGA 616
Db 436 GATCTTTAGATTTGGAGACATCCATAGAGATTTTACATCCGAGAGATGAAAGAA 495
QY 617 AACGTGATGTGTGCTGACTGACTGAAGAACTTGAATGTTCTCTGAACTGTAAATACAG 676
Db 496 AACACAGTGAATTTTGGGGCAAGGAATGCAATGTGAAGTGAAGTGAAGTGAATAGT 555
QY 677 TGGATCAAAAAGGTGAAAAAAGACATTAAGAGAGCTAAATGATGATCATAGT 736
Db 556 CTTTGGAGATGGGGAATGAGCCAGCATACAGTCCACTCAAAATTAAGACACTCC 615
QY 737 AGTCGTCACATACATATTTAGATGATGTTGTAAGCTTGAAGCCGAGACAGAAAT-----GAT 790
Db 616 AGCAGATCAGATGAATTTTACATACAGATTTGTCAATGTAATTAATTAAGAGAGCA 675
QY 791 CCCCAATTCAGAGAACTGTGATGAGCTGTCAATGTAATCTCAATTTGATAGT 850
Db 676 GCTGAAGATGATCATGATGATTTCCCTCGGCATATTTCTCAAAAGTTCCACTTTGTGAT 735
QY 851 CTTGCTGCAAGTGAAGAGCAAGCCAACTGAGCTGAAGTGTGAGACTTAAGAGAGC 910
Db 736 TTGGCAGAGATCAGAAAGATTAACCAAAACGAGGAAATCTGTGACGTTCAAGAAATCC 795
QY 911 TGCACATCAACGCGACCTGTTTATCCCTGACAGGTTATTAAGAACTTAGGAGC-- 968
Db 796 ATTCAATCAATAGTATGATGCTGCTTTAGGAATTAATTAAGCGCTCTGGAGACCA 855
QY 969 -GCCAGCTGTGATTTAATTAACACTACAGAGACGAACTCACAGAAATTTCCAAAT 1027
Db 856 CGCAGGAAGATTCATATTAATTCATATTAAGGATGCTAAATTAACCGGCTTCTGAAGAT 915
QY 1028 TCATTGGAGAGAAATGCTAAACGTTATTAATTTGACAAATTAAGCCAGTTTC-----T 1081
Db 916 TCCTGGAGAGAGATGCTAAGATGTCATGATCAGATGTGACAGCCCTCCTCGAAT 975
QY 1082 TTTGATGAGACTCTAAATACACTTTCAGTTGCCAGTACTGCAAAACATGTGAGAAATACT 1141
Db 976 TTTGATGAGTCTTAATTTCTCTCAATATATGCCAAGAGACGAGAACTTAGAAACAA 1035
QY 1142 CCCCATGTTAA 1152
Db 1036 CCCACTGTAA 1046

```

```

RESULT 3
US-09-723-096-1
; Sequence 1, Application us/09723096
; Patent No. 6448026
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6448026el motor proteins and methods for

```


QY	275	TTTCATTTGATGCTGTGATATTTAAATTCACGAAATCAACACTCAAAATTTACCAAGAAATA	33
Db	130	TTCACTTTTGTATTTGTTTTTGGCAAAAATTCACCTCAAGATGAAATTTAAACACATGT	189
QY	335	GCAGTACCTATCATACGATCAAGCTTTGCAGGAGTATATATGGCAAAATATTTGGCATACGGA	399
Db	190	ATYAAAGCCCCATAGTGTGTGCACCTCATYGGAGGGCATATATGAACACTGTTTTTGGCCCTATGGA	249
QY	395	CAGACATCTTTCAGGCAAGACGTACACATGATGGGAACACCAATTTATTT	444
Db	250	CAAACTGGATTTGGGAAGACATACACCATTTGGAGGGGCCATATTTGCTTCAGTTGTGGAG	309
QY	445	-----GGGCATTAATACCCCAAGCCATACAGGAAGTTTTTAAATTAATTCAGAGATA	496
Db	310	GGCCAAAAGGCTATCAATCTCCGAGCATTCATCAAAAATATTTCAAAACACTCTCGAACT	369
QY	497	CCGAACAGAGGTTCTTCTTAAGAGTTCTTAATTTGAGAGATTTCCATCTGAAACCTGAAA	556
Db	370	CTTAGCATTTGCTTTAATGTAAAGATATCTTAATTAAGAAAGTACAGGAAGACCTATAGA	423
QY	557	GACCTACTGTGTGATGACAGAAAGAAAGAGCCCTTGGAAATTCGCGAGAGATTTTAATAGA	618
Db	430	GATCTTCTAGAAATTTGAGAGACATCCATGAAGATCTTCACATCCGAGAAAGATGAAAAAGGA	489
QY	617	AACGTGATGTGCTGCTGACCTGACTGAAAGAACTGTAAATGTTCTCTGAACATGTAAATACAG	676
Db	490	AACACAGATGATTTGTTGGGGCCAAAGAAATGCCATCTGGAGAGTGGCAGGTGGAAGTGTAGT	549
QY	677	TGGATCAAAAAGGGTGGAAAAAAGACATTAATGAGAGAGACTAAATGAATGATCATAGT	736
Db	550	CTTTTGGAGATGGGGAATGCAACGCCAGACATACAGTATCCACTCAATATGAATGAGACCTCC	609
QY	737	AGTCGTTACATCAATATTTAGATGATGTTTGAAGACCGAGACAGAAAT-----GAT	790


```

OY 395 CAGACATCTTCAGCAGACGATACACAATGATGGNAACCAATTCATT----- 444
DB 250 CAAATCGATCTGGGAAAGCATACACATTCGAGGGGGCCATATTCCTTCAGTTGGAG 309
OY 445 -----GGGCAATAATCCCCAACCCATACAGAAATTTTAAATTTATTCAGAGATA 496
DB 310 GGGCAAAAGGATATTCCTCGAGCTATTCAGAAATTTTCAAAAGCATCTCTGMAAT 369
OY 497 CCGAAGAGAGATTTCTTATAGATTTCTTATAGAGATTATACAAATTAACCTGGA 556
DB 370 CCTACATGATCTTAATGATAAGTATCTTATAGAGATTATACAAAGAACCTTAAGA 429
OY 557 GACCTACTGTGTGATGACAGAAAGAACCCCTTGGAATTCGAGAGATTTTATAGA 616
DB 430 GATCTCTAGATTTGGAGCATCCATGAAAGATCTTCACATCCGAGAGATGAAAAAGA 489
OY 617 AACGTGATGTGCTGACCTGACGAGAACTTGTAAAGTCTGTAAACATGTAATACAG 676
DB 490 AACACAGTATGTGGGGCCAAAGAAATCCATGTGAGAGTGCAGGTGAATGATGAGT 549
OY 677 TGGATCAAAAAGGTGAAAGAACACATTTGAGAGACATAAATGATATCATAGT 736
DB 550 CTTTGGAGATGGGAAATGACACGACATACAGGTACACCTCAATGATGAGACTCC 609
OY 737 AGTCCTCATACATATATTTAGATGATTTGAAAGCCGAGAGAGAAAT-----GAT 790
DB 610 AGCAATCATACATGCAATTTTACATACAGCATTTGTCAAGTTCAATAAATATGAGAGA 669
OY 791 CCCACAAATTCAGAGAACTGTGAGAGCTGATGATCTGATCTGACTGAAATTTGGTAGAT 850
DB 670 GCTGAGATGTGATCATGTGATTCCTCGGCAATTTGTCTCAAGTTCACACTTGTGAGAT 729
OY 851 CTGTGTGAGAGAAAGCAAGCAAACTGAGAGCTGAAGTGTGAGACTTAAGAGAGC 910
DB 730 TTGGAGAGATGAGAAAGTAAACCAAAAGGGGAAATACTGTGAAAGGTTTCAAAANTCC 789
OY 911 TGCACATCAACCCGAGCTGTTTATCTTGTGACAGGTTATTAAGAGCTTATGACAGC-- 968
DB 790 ATTCAATCAATAGTGAATGTGCTTGAAGAAATGTAATAGCGCTTGTGGGACCA 849
OY 969 -GCCAGCTGTGTATTAATAACTACAGAGACAGCAAACTGACAGAAATTCCTCAAAAT 1027
DB 850 CGCAGAGAGATGTCATATTCATATAGGAGTGTCTAAATTAACCCGGCTCTGAAGAT 909
OY 1028 TCATGGAGAGAAATGCTAAACGTTTAAATTTGCACATTAACCCAGTTTC-----T 1081
DB 910 TCTCTGGAGGAGTGTGACAGCTGTCAATGATCAGATGTGAGCCCTCCTCCTCGAT 969
OY 1082 TTTGATGAGACTCTTAAGTACATTCAGTTTGGCAGTACTGCCAAACATGTGAGA 1135
DB 970 TTTGATGAGTCTTAATTTCTCAAAATATGCAACAGAGCAGGAACATTAGA 1023

```

```

RESULT 7
US-09-592-054-5
: Sequence 5, Application us/09592054
: Patent No. 6440684
: GENERAL INFORMATION:
: APPLICANT: Berand, Christophe
: APPLICANT: Finer, Jeffrey
: APPLICANT: Sakowicz, Roman
: APPLICANT: Wood, Kenneth
: TITLE OF INVENTION: No. 6440684el motor proteins and methods for
: TITLE OF INVENTION: their use
: FILE REFERENCE: 1016
: CURRENT APPLICATION NUMBER: us/09/592, 054
: CURRENT FILING DATE: 2000-07-20
: NUMBER OF SEQ ID NOS: 8
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 5
: LENGTH: 1421
: TYPE: DNA

```

```

: ORGANISM: Human
US-09-592-054-5

```

```

Query Match 1.4%; Score 135.2; DB 4; Length 1421;
Best Local Similarity 50.6%; Pred. No. 9,8e-24;
Matches 468; Conservative 0; Mismatches 408; Indels 48; Gaps 4;

```

```

OY 257 GTTGATGGAGCAAAAGCTTTCATTTGATCGATCGTGTAAATTCACAGAACCAACAGT 316
DB 123 GTTGGTACAGATTAATCTTTCACCTACGATTTTGTATTTGATCCCTCTCTGAAACGAA 182
OY 317 CAAATTTACCAAGAAATATACAGTACCTATACATACGATTCAGTTTGGAGGATATATGCG 376
DB 183 GAAGTCTCAATATACAGACTAGCGCACATCAATAAAGGTATTTAAAGATATATATGA 242
OY 377 ACAATTTTTCATACGAGACAGACATCTTACGCAAGACCTACACATATGATGGAAACCA 436
DB 243 ACGGTCTGGCTATGGGACAGCTGGCTGTGAAAAAACCCTATTCATATGGAGGTGCTAT 302
OY 437 AATTCATTTG-----GGCATTAATACCCCAAGCCATACAGAAAT 475
DB 303 ACTGAGAGCAAGACATACCAACAGTGGGTTATTTCTTAGGTTAATACACTGCTC 362
OY 476 TTTAAATTTATTCAGAGATACCGAACAGAGATTTCTTAAGATTTCTTAATGAG 535
DB 363 TTCAAAAGAAATGATTAATAAGAGTACTTTGAATTTACTGAAAGTGTCTTAAGTAG 422
OY 536 ATTACATATGAATCTGTGAAGAACCTACTGTGTGATGACAGAAAGAAAGCCCTTGGCA 595
DB 423 ATTTCAATGAAGAAATTTTGTGATCTTATCCATCCATCTCGAGAAACCTCAATTAAT 482
OY 596 ATTGGAGAGATTTTAAATGAAGCGTATTTGTGTCAGCTGACAGAACTGTAATG 655
DB 483 ATACGAGAGATCTTAAGAGAGCATTAAGATTTGGAGCTACTGAGAAAGCTGTTTGG 542
OY 656 GTTCTGAAATGTAATACATGATGATCAAAAGGTTGAAAAAAGACATTTAGAGAG 715
DB 543 GTTGCTGTGATACGTGTTTCTGTTTGAACAGGGCAACCTAGAGACTGTGGGCTCC 602
OY 716 ACTAAATATGATCATATGATGCTGTTGATACATATTTAGAAATGATTTTGAAGC 775
DB 603 ACGGTATGTAACCTGAGTCCGAGTCTCCGATCTCATGTCATCTTACAT--CTCCTTAGAGC 660
OY 776 CGAGACAGAAATGATCCCAATTCAGAGAACTGTGATGAGAGTGTGATGATCTGAC 835
DB 661 AAAGAAAGAAAGTACAGAAATAGCAG-----CTTTGCTCCAG 701
OY 836 TTGAATTTGATGATCTTGTGTCAGTGAAGAGCAAGCAAACTGAGAGCTGAGTGTG 895
DB 702 CTGCAATCTTGTAGACTGCTGCTGATCAGAAAGACAGAAAGCAAGGCTGAAGGGAT 761
OY 896 AGACTTAAGAGAGCTGCAACATCAACCGCAGCTGTTTATCTTGGAGAGGTTTAAAG 955
DB 762 CECTTAAGAGAGGATTAATTAATTAACCGAGCCTCTATGCTTGGAAATTAATCAAT 821
OY 956 AAGCTTAGGAGCGGCGAGCTGTGATTTAATTAATCTACAGAGACAGCAAACTGCCAGA 1015
DB 822 GCTCTTGGAGATGACAAAAAGGCTGTGTTGTGCTTACAGAGATTTCCAAAGTTGACTGA 881
OY 1016 ATTCTCAAAATTCATTTGGAGAAATGCTAAAGGTTAATTAATTTGCAATTAAGCCA 1075
DB 882 CTGCTTCAAGATTTCTAGAGAGTAAATAGCCATCTATGATAGCGTGTGAGTCT 941
OY 1076 GTT-----TCTTTGATGAGAGCTTAAGTACATCTCAATTTGCCAGTACTGCCAAACT 1129
DB 942 GCTGACTCCAACTAGAGAAACATTAATACCTTGTGCTAGTGCAGAGACAGCAAGAAA 1001
OY 1130 GTGAGAAATATCTCCCATGTTAAT 1153
DB 1002 ATCAAGAACAACTATTTGTTAAT 1025

```

```

RESULT 8

```

US-09-592-054-3
 ; Sequence 3, Application US/09592054
 ; Patent No. 6440684
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; APPLICANT: Beraud, Jeffrey
 ; APPLICANT: Sakowicz, Roman
 ; APPLICANT: Wood, Kenneth
 ; TITLE OF INVENTION: No. 6440684e1 motor proteins and methods for
 ; FILE OF INVENTION: their use
 ; FILE REFERENCE: 1016
 ; CURRENT APPLICATION NUMBER: US/09/592,054
 ; CURRENT FILING DATE: 2000-07-20
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 1569
 ; TYPE: DNA
 ; ORGANISM: Human
 ; US-09-592-054-3

Query Match 1.48; Score 135.2; DB 4; Length 1569;
 Best Local Similarity 50.68; Pred. No. 1e-23;
 Matches 468; Conservative 0; Mismatches 408; Indels 48; Gaps 4;

QY 257 GTTGATGGAGCAAAAGCTCTTCAATTCGATCGTATTTATTCACGAATCAACAAGT 316
 DB 187 GTTGATAGAGTAATCTTCCACACGATTTTGTATTTGATCCCTACGAAACAGAA 246
 QY 317 CAAATTTACCAAGAAATAGCAGTACCTATCATAGCATGACTTTGACGGATATATGCG 376
 DB 247 GAAGCTTCAATACAGCAGTACCCACTATAAAGGTATTAAGGTATATAATCA 306
 QY 377 ACAATATTGCTATAGGAGACATCTTCAGCAGAGCTACACATATGATGGAACCA 436
 DB 307 ACGGTCTGGCTATGGCAGACGCTGCTGTGAAAAACCTATCATATGGAGGTGCATAT 366
 QY 437 AATTCATTG-----GGCATATAACCCCAAGCCATACAGAACTT 475
 DB 367 ACTGACAGACAGAGAAATGACCAACAGTGGGGTATTCCTAGAGGTATACAACTGCTC 426
 QY 476 TTTAAATATTTCAGAGATACCGAAGAGAGTCTTCTTAAGAGTTCTTATATGAG 535
 DB 427 TTCAAGAAATGATAAAAAGAGTCTTGAATTTACTGTAAGAGTCTTACTTAGAG 486
 QY 536 ATTACAAATGAAGTGTGAAGACCTACTGTGTGATGACAGAAAGAAAGCCCTTGAA 595
 DB 487 ATTACAAATGAAGAAATTTTGTGATCTGTATGCCATCTCGAAGAAAGCTCAATTAAT 546
 QY 596 ATTGCGAGATTTTATATGAAGAGTATGTTGCTGACCTGACCTGAAAGAACTTATATG 655
 DB 547 ATACGAGAGATCTTAAGAGAGCATAAAGATTGGGACCTCAGAGAAAGACCTTTTG 606
 QY 656 GTTCCGAACTGATATACAGTGTGATCAAAAAGGGTGAAGAAAGACATATATGAGAG 715
 DB 607 GTTGCTTGGATCTGTCTTCTGTGTGAAACAGGCAACACTCTAGAGCTGTGCTTCC 666
 QY 716 ACTAAATATGATATCATAGTATGCTTCCATCATATATTTAGATGATTTTGAAGC 775
 DB 667 ACGGCTATGATCCAGAGCTGCCGATCTCATGCAATTTTAACT--CTCCTTAGAGC 724
 QY 776 CGAGACAGAAATGATCCCAAAATTCAGAAAGTGTATGAGAGCTGTATGATCTAC 835
 DB 725 AAAGAAAGAAAGTACAGAAATAGAG-----CTTTGCTTCAAG 765
 QY 836 TTGAATTTGATGATCTGTGCAAGTGAAGAGCAAGCAACTGAGCTGAGGTG 895
 DB 766 CTGACATCTGTAGCTCTCTGTGATGAGAAAGCAAGAAAGCAAGGCTTAAGGGAT 825
 QY 896 AGACTTTAAGGAAGGCTGCAACATCAACCCAGCTGTTTATCTTGTGACAGGTTATTAAG 955
 DB 826 CGCTTAAAGAGGATTTATATTAATTAACGAGGCTCTATGTTGGAAATGTATTAAGT 885

QY 956 AAGCTTAGCAGGAGCCAGCGCTGTGATTTATTAATTAACATACAGAGACGAACATCAACGA 1015
 DB 886 GCCTCTTGAGATGACAAAAGAGGTGGCTTTGTGCCCTACAGAGATCCAAAGTTGACTGA 945
 QY 1016 ATTTCGCAAAATTCATTGGAGAGAAATGCTAAAGCGTATATTTGACAAATTCAGGCA 1075
 DB 946 CTGCTTCAAGATCTCTAGAGAGTAAAGCAATACCTTATATATAGCTGTGTGAGTCT 1005
 QY 1076 GTT-----TCTTTGATGAGACTTAAGTACACTTCACTTGTGCGAGTACTGCCAAAGAT 1129
 DB 1006 GCTGACTCCAAATCTAGAGAGAAATTAATTAATACCTTCTGCTATGCGAGACAGCAAGAAA 1065
 QY 1130 GTGAGAAATACCTCCCATTTAT 1153
 DB 1066 ATCAGAAACCAACCTATTTAT 1089

RESULT 9
 US-09-592-054-7
 ; Sequence 7, Application US/09592054
 ; Patent No. 6440684
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; APPLICANT: Beraud, Jeffrey
 ; APPLICANT: Sakowicz, Roman
 ; APPLICANT: Wood, Kenneth
 ; TITLE OF INVENTION: No. 6440684e1 motor proteins and methods for
 ; FILE OF INVENTION: their use
 ; FILE REFERENCE: 1016
 ; CURRENT APPLICATION NUMBER: US/09/592,054
 ; CURRENT FILING DATE: 2000-07-20
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 7
 ; LENGTH: 4127
 ; TYPE: DNA
 ; ORGANISM: Human
 ; US-09-592-054-7

Query Match 1.38; Score 127.2; DB 4; Length 4127;
 Best Local Similarity 50.18; Pred. No. 1.6e-21;
 Matches 463; Conservative 0; Mismatches 413; Indels 48; Gaps 4;

QY 257 GTTGATGGAGCAAAAGCTCTTCAATTCGATCGTATTTATTCACGAATCAACAAGT 316
 DB 229 GTTGATCTGATTAATCTTCCACACGATTTTGTATTTGATCCCTACGAGAGAGAA 288
 QY 317 CAAATTTACCAAGAAATAGCAGTACCTATCATAGCATGACCTTTGACGGATATATGCG 376
 DB 289 GAAGTCTTCAATTAAGAGAGTACCCGCTCATATAAAGCATATTTAAAGATATTAATGCA 348
 QY 377 ACAATATTGCTATAGGAGACATCTTCAAGGCAAGACGTACACAAATGATGGAACACCA 436
 DB 349 ACGGTCTGGCTTATGGCAGACGCTGTGAAAAACCTATTAACGAGAGGTGATAC 408
 QY 437 AATT-----CATGGCAATAATACCCAAAGCCATACAGAAAGT 475
 DB 409 ACTGCGAGACAGAGATGAAACCAACAGTGGCATTTTCTTAGGCTATATCAAAAGCTC 468
 QY 476 TTTAAATATTTCAGAGATACCGAAGAGAGTCTTCTTAAGAGTTCTTATATAGAG 535
 DB 469 TTCAAGAAATGATCAAAAGAGTCTTGAATTTACTGTGAAGAGTGTCTTACTTAGAG 528
 QY 536 ATTACAAATGAAGTGTGAAGACCTACTGTGTGATGACAGAAAGAAAGCCCTTGAA 595
 DB 529 ATTACAAATGAAGAAATTTTGTGATCTTCTATGCCATCTGTGAGAAAGCTCAATTAAT 588
 QY 596 ATTGCGAGATTTTATATGAAGAGTGTATGTTCTGACCTGACTGAGAAAGTGTATG 655
 DB 589 ATAGGGAGAGTCTTAAGAGAGCATAAAGATTGTGGAGCTCACTGAGAAAGAGCTTTTA 648
 QY 656 GTTCTGAACATGTATTAATCACTGTGATCAAAAAGGGTGAAGAAAGACAGATTTATGAGAG 715

```

Db 649 GTTGCCCTGGATCTGTTCTCTGTTGGACGAGGCGACACATCTAGACTGTGGCTCC 708
QY 716 ACTAAATGATGATCTAGTACGTTCACATACATATTTAGATATGTTGAAAGC 775
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 709 ACAGCTATGAGTCCGAGTGTCCGATCTCATGCCATCTTTACAAT--CTCCTTAGAGC 766
QY 776 CGAGACGAATATATCCCAATTTACAGAACTGTGATGAGAGCTGTATGTCTCAC 835
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 767 AAGGAAAGAAAGTACAGAAATAGCAG-----CTTTCCTCCAG 807
QY 836 TTGAATTTGGTATCTGTGCGAGTGAAGAGCAAGCAAACTGAGTGAAGGTG 895
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 808 CTCATCTTGTATACCTCGGTGATCAGAAAGACAGAAACCAAGCTGAAGGGAT 867
QY 896 AGACTTAAGAGAGCTGCAATCAACCCGAGCTTGTATCTTCCGTGAGACAGTTATTAAG 955
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 868 CGCTTAAGAGGGTATTAATATTAACCGAGGCTCCTATGCTTGGAAATGTAATAGT 927
QY 956 AAGCTTAGCAGCGCCAGGCTGTGATTTATAACTACAGAGACAGCAAACTCACCA 1015
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 928 GCTCTTGAGANTACAAAAGGCTAGCTTGTGCCCTACAGAGATTCCAAAGTTAAGTCGA 987
QY 1016 ATTCTCCAAATTCATTTGGAGAAATGCTAAACGGTTATTTGCAATTAAGCCA 1075
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 988 CTGCTGAGATCTCTAGAGGTAAACAGCCACACTCTTAAGTAGAGCTGTGAGTCT 1047
QY 1076 GTT-----TCTTTGATGAGACTTAAGTACACTTCAGTTGCGCACTGCAATTAAG 1129
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1048 GCTGACTCCAACTAGAGAAACATTAAGTACCTTCGCTATGCTGACAGCAAGAAAA 1107
QY 1130 GTGAAATATCTCCCATGTTAAT 1153
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1108 ATCAGAACAAACCTATTGTTAAT 1131
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 10
US-09-592-054-1
; Sequence 1, Application US/09592054
; Patent No. 6440684
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Finer, Jeffrey
; APPLICANT: Sakowicz, Roman
; APPLICANT: Wood, Kenneth
; TITLE OF INVENTION: No. 6440684e1 motor proteins and methods for
; FILE REFERENCE: 1016
; CURRENT APPLICATION NUMBER: US/09/592,054
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 4308
; TYPE: DNA
; ORGANISM: Human
US-09-592-054-1

```

```

Query Match 1.2% Score 116; DB 4; Length 4308;
Best Local Similarity 49.4%; Pred. No. 1e-18;
Matches 456; Conservative 0; Mismatches 420; Indels 48; Gaps 4;

```

```

QY 257 GTTGATGGACAAAGTCTTTCATTTGATCGTATTTAATTTCTCAGAAATCAGCAAGT 316
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 187 GTTGGTACAGATTAATCTTACCTACGATTTGTATTTGATCCCTACTAGAACAGAA 246
QY 317 CAAATTTACCAAGAAATAGCAGTATCATATCATGATCAGCTTTGCAAGGATTAATGCC 376
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 247 GAAGTCTTAATATACAGAGTAAAGCCATCAATAAAGGTATTAAGAGATTAATGCA 306
QY 377 ACAATATTTGATAGCAGACATCTTACGAGCAAGCAGTACAGTAAAGTGAAGACCA 436
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 307 ACGGTCTGCGCTATGAGGAGACTGGCTGTGAAAAAACCTATTCAATGGAGGTGATAT 366

```

```

QY 437 AATTCATG-----GCCATATACCCCAAGCCATACAGAAATG 475
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 367 ACTGACAGACAGAAATGAAACCAACAGTTGGGGTTATTCCTAGAGGTAATCAACTGCTC 426
QY 476 TTTAAATTTATGAGGATACGGAACAGAGAGTTCTTCTAGAGTTCTTATATGAG 535
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 427 TTCAAAGAAATGTAAAGAGTGAATTTGATTTACTGTGAAGGTGTACTTAGAG 486
QY 536 ATTTCATGAACTGTGAAGACCTTCTGTGATGACAGAAAGAAAGACCTTTGAA 595
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 487 ATTACATGAAAGAAATTTTGATCTCTATGTCATCTGTGAGAAAGCTCAATTAGAG 546
QY 596 ATTGCGAGAGTTTAAATAGAAACGTGTATGTGCTGACCTGACTGAAGACTTGAATG 655
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 547 ATTACATGAAAGAAATTTTGATCTCTATGTCATCTGTGAGAAAGCTCAAGTTTG 606
QY 656 GTTCCCTGACATGTAATACAGTGAATCAAAAGGTTGAAAAAACACATTAATGAGAG 715
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 607 GTTGCCCTGGATCTGTTCTCTGTTGGAAACAGGCAACCTTAGACTGTGGCTCC 666
QY 716 ACTAAATGATGATCTAGTACGTTCACATACATATTTAGATATGTTGAAAGC 775
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 667 ACGGCTATGAACTCCAGTGTCCGATCTCATGCCATCTTAACAAT--CTCCTTAGAGC 724
QY 776 CGAGACGAATATATCCCAATTTACAGAACTGTGATGAGAGCTGTATGTTCTCAC 835
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 725 AAGGAAAGAAAGTACAGAAATAGCAG-----CTTTCCTCCAG 765
QY 836 TTGAATTTGGTATCTGTGCGAGTGAAGAGCAAGCAAACTGAGTGAAGGTG 895
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 766 CTGATCTTGTAGACCTCGGTGATCAGAAAGACAGAAACCAAGCTGAAGGGAT 825
QY 896 AGACTTAAGAGAGCTGCAATCAACCCGAGCTTGTATTTCCGTGAGACAGTTATTAAG 955
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 826 CGCTTAAGAGGGTATTAATATTAACCGAGGCTCCTATGCTTGGAAATGTAATAGT 885
QY 956 AAGCTTAGCAGCGCCAGGCTGTGATTTATAACTACAGAGACAGCAAACTCACCA 1015
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 886 GCTCTTGAGANTACAAAAGGCTAGCTTGTGCCCTACAGAGATTCCAAAGTTGACTCGA 945
QY 1016 ATTCTCCAAATTCATTTGGAGAAATGCTAAACGGTTATTTGCAATTAAGCCA 1075
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 946 CTGCTGAGATCTCTAGAGGTAAACAGCCACACTCTTAAGTAGAGCTGTGAGTCT 1005
QY 1076 GTT-----TCTTTGATGAGACTTAAGTACACTTCAGTTGCGCACTGCAATTAAG 1129
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1006 GCTGACTCCAACTAGAGAAACATTAATACCTTCGCTATGCTGACAGCAAGAAAA 1065
QY 1130 GTGAAATATCTCCCATGTTAAT 1153
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1066 ATCAGAACAAACCTATTGTTAAT 1089
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 11
US-09-641-806-3
; Sequence 3, Application US/09641806
; Patent No. 6395527
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. 6395527e1 motor proteins and methods for
; FILE REFERENCE: 1034
; CURRENT APPLICATION NUMBER: US/09/641,806
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1011
; TYPE: DNA
; ORGANISM: Human
US-09-641-806-3

```

Query Match 1.0%; Score 98; DB 4; Length 1011;
 Best Local Similarity 48.3%; Pred. No. 1.3e-14;
 Matches 346; Conservative 0; Mismatches 355; Indels 15; Gaps 2;

```

OY 359 TTGCAGGATATTAATGAGCAATATTTGCATACGAGACAGACATCTTCAGGCAAGACGTAC 418
   || || || || || || || || || || || || || || || || || || || || || || ||
Db 214 TTGCAGGCTTCATGCTCCACTGTCTTGTGCTATGCTACAGCGGGCTCAGGAAACATAC 273
OY 419 ACAATGATGGAAACACCA-----ATTCAITGGGCATTAATACCCCAAGCCATA 466
   || || || || || || || || || || || || || || || || || || || || || || ||
Db 274 ACCATGGGGAGGCGCCACCTCCCTCTTGAGATGAGCAGGCGCATGTCCGAGGGCCATG 333
OY 467 CAGGAATTTTAAATTTTACAGAGATCCGAAACAGAGATTTCTTCTTAAGATTTCT 526
   || || || || || || || || || || || || || || || || || || || || || || ||
Db 334 GCGAGGCGCTTCAGCTCATGATGAGAACGACCTGCTTGACTGTGTATCATGTGTCC 393
OY 527 TATATGAGATTTACATTAATGAACTGTGAAAGACTACTGTGTGATGACAGAAAGAAAG 586
   || || || || || || || || || || || || || || || || || || || || || || ||
Db 394 TACCTGGAAGTGTACAGAGAGAGATTCGAGACCTGCTGAGAGTGGGCTGCGCAGCCGT 453
OY 587 CCCTTGGAAATTCGCGAGGATTTTAAAGAAAGGTATGTTGCTGACCTGACGTAGAGAA 646
   || || || || || || || || || || || || || || || || || || || || || || ||
Db 454 GACATCCAGCTCCGGGAAGATGAGCGCGGAATGTTGTGCTGTGGGGGTGAAGAGGTC 513
OY 647 CTGTAAATGTTCTGAACATGTATACAGTGTATACAAAGGGTGAAGAAACAGACAT 706
   || || || || || || || || || || || || || || || || || || || || || || ||
Db 514 GACGTGAGAGGCTCGATGAGTGTGCTGAGCTCTGAGATGGGCAACGCGCGCGGCAC 573
OY 707 TATGAGAGACTTAAATGATGATCATAGTGTGCTGATACATATTAATTAAGATGAT 766
   || || || || || || || || || || || || || || || || || || || || || || ||
Db 574 ACGGAGACGACGACCTCAACACCTGTCTAGCGGCTCAGACAGGTGTTACCGTGACC 633
OY 767 GTTGAACCCGAGACAGAAATGATCCCAAAATTCAGAACTGTGATGAGCTGTCAATG 826
   || || || || || || || || || || || || || || || || || || || || || || ||
Db 634 CTGGAGCAGCGGGGGCGGCCCAAGCCGCTACCCCGCCCGCCGCGGCGACGTGCTC 693
OY 827 GTATCTCACTGGAATTTGTGATGATCTTCTGAGTGAAGAAAGAGCAAGCAACTGAGCT 886
   || || || || || || || || || || || || || || || || || || || || || || ||
Db 694 GTCTCCAGTTCACACTGTGTGAGCTGCGGGCTCAGAGAGGGTCTCAACGCGGCGAC 753
OY 887 GAAGGTGTGAGACTTAAAGAGGCTGCACATCAACCGCAGCTGTTATCTTGGACAG 946
   || || || || || || || || || || || || || || || || || || || || || || ||
Db 754 ACCGGCGAGCGGCTCAAGAGAGATCCAGATCAACAGACGCTCTGGGCGCTGGCAGC 813
OY 947 GTTATTAGAAGCTTAAGAGCGGCGCAGGCTGTGATTT---ATAACTACAGAGACAGC 1003
   || || || || || || || || || || || || || || || || || || || || || || ||
Db 814 GTCATCAGCGCCCTGGGGGACCTCAGCGCGCGGCGGAGCCACATACCTACCGGACTCC 873
OY 1004 AAACCTACAGAAATTTCCAAATTCATTTGGAGAAATGCTAAACGGTTATAT 1059
   || || || || || || || || || || || || || || || || || || || || || || ||
Db 874 AAGATCACCGGATCTCAAAAGACTCGCTGGGCGGGAACCGCAAGCGGTGATGAT 929

```

RESULT 12
 US-09-641-806-1
 ; Sequence 1, Application US/09641806
 ; Patent No. 6395527
 ; GENERAL INFORMATION:
 ; APPLICANT: Berand, Christophe
 ; APPLICANT: Freedman, Richard
 ; TITLE OF INVENTION: No. 6395527el motor proteins and methods for
 ; TITLE OF INVENTION: their use
 ; FILE REFERENCE: 1034
 ; CURRENT APPLICATION NUMBER: US/09/641, 806
 ; CURRENT FILING DATE: 2000-08-17
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 1026
 ; TYPE: DNA
 ; ORGANISM: Human
 ; US-09-641-806-1

Query Match 1.0%; Score 98; DB 4; Length 1026;
 Best Local Similarity 48.3%; Pred. No. 1.3e-14;
 Matches 346; Conservative 0; Mismatches 355; Indels 15; Gaps 2;

```

OY 359 TTGCAGGATATTAATGAGCAATATTTGCATACGAGACAGACATCTTCAGGCAAGACGTAC 418
   || || || || || || || || || || || || || || || || || || || || || || ||
Db 217 TTGCAGGCTTCATGCTCCACTGTCTTGTGCTATGCTACAGCGGGCTCAGGAAACATAC 276
OY 419 ACAATGATGGAAACACCA-----ATTCAITGGGCATTAATACCCCAAGCCATA 466
   || || || || || || || || || || || || || || || || || || || || || || ||
Db 277 ACCATGGGGAGGCGCCACCTCCCTCTTGAGATGAGCAGGCGCATGTGCCAGGGCCATG 336
OY 467 CAGGAATTTTAAATTTTACAGAGATCCGAAACAGAGATTTCTTCTTAAGATTTCT 526
   || || || || || || || || || || || || || || || || || || || || || || ||
Db 337 GCCAGGCGCTTCAGCTCATGATGAGAGACAGCTGCTTGACTGTGTGTATCATGTGTCC 396
OY 527 TATATGAGATTTACATTAATGAACTGTGAAAGACTACTGTGTGATGACAGAAAGAAAG 586
   || || || || || || || || || || || || || || || || || || || || || || ||
Db 397 TACCTGGAAGTGTACAGAGAGAGATTCGAGACCTGCTCGAGAGTGGGCTACGCCGCT 456
OY 587 CCCTTGGAAATTCGCGAGGATTTTAAAGAAAGGTATGTTGCTGACCTGACGTAGAGAA 646
   || || || || || || || || || || || || || || || || || || || || || || ||
Db 457 GACATCCAGCTCCGGGAAGATGAGCGCGGAATGTTGTGCTGTGCGGGGTGAAGAGGTC 516
OY 647 CTGTAAATGTTCTCTGAACATGTATACAGTGTATACAAAGGGTGAAGAAACAGACAT 706
   || || || || || || || || || || || || || || || || || || || || || || ||
Db 517 GACGTGAGAGGCTCGATGAGTGTGCTGAGCTCTGAGATGGGCAACGCGCGCGGCAC 576
OY 707 TATGAGAGACTTAAATGATGATCATAGTGTGCTGATACATATTAATTAAGATGAT 766
   || || || || || || || || || || || || || || || || || || || || || || ||
Db 577 ACGGAGCCAGCAGCCTCAACACCTGTCTAGCGGCTCAGACAGGCTCTCACCGTGACC 636
OY 767 GTTGAACCCGAGACAGAAATGATCCCAAAATTCAGAACTGTGATGAGCTGTCAATG 826
   || || || || || || || || || || || || || || || || || || || || || || ||
Db 637 CTGGAGACGCGGGGGCGGCCCAAGCGCTACCCCGCCCGCGGCGGCGACGTGCTC 696
OY 827 GTATCTCACTGGAATTTGTGATGATCTTCTGAGTGAAGAAAGAGCAAGCAACTGAGCT 886
   || || || || || || || || || || || || || || || || || || || || || || ||
Db 697 GTCTCCAGTTCACACTGTGTGAGCTGCGGGCTCAGAGAGGGTGTCTCAAGAGGCGCAG 756
OY 887 GAAGGTGTGAGACTTAAAGAGGCTGCACATCAACCGCAGCTGTTATCTTGGACAG 946
   || || || || || || || || || || || || || || || || || || || || || || ||
Db 757 ACCGGCGAGCGGCTCAAGAGAGATCCAGATCAACAGAGACCTCTGGGCGCTGGCAGC 816
OY 947 GTTATTAGAAGCTTAAGAGCGGCGCAGGCTGTGATTT---ATAACTACAGAGACAGC 1003
   || || || || || || || || || || || || || || || || || || || || || || ||
Db 817 GTCATCAGCGCCCTGGGGGACCTCAGCGCGCGGCGGAGCCACATACCTACCGGACTCC 876
OY 1004 AAACCTACAGAAATTTCCAAATTCATTTGGAGAAATGCTAAACGGTTATAT 1059
   || || || || || || || || || || || || || || || || || || || || || || ||
Db 877 AAGATCACCGGATCTCAAAAGACTCGCTGGGCGGGAACCGCAAGCGGTGATGAT 932

```

RESULT 13
 US-09-724-519-5
 ; Sequence 5, Application US/09724519
 ; Patent No. 6414121
 ; GENERAL INFORMATION:
 ; APPLICANT: Wood, Kenneth
 ; APPLICANT: Finer, Jeffrey
 ; APPLICANT: Berand, Christophe
 ; APPLICANT: Mak, John
 ; TITLE OF INVENTION: Methods of screening for modulators of
 ; TITLE OF INVENTION: cell proliferation and methods of diagnosing cell
 ; TITLE OF INVENTION: proliferation states
 ; FILE REFERENCE: 1014A
 ; CURRENT APPLICATION NUMBER: US/09/724, 519
 ; CURRENT FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 09/592, 037
 ; PRIOR FILING DATE: 2000-06-12
 ; PRIOR APPLICATION NUMBER: 09/428, 156
 ; PRIOR FILING DATE: 1999-10-27

```

: NUMBER OF SEQ ID NOS: 10
: SOFTWARE: FASTSQ for Windows Version 4.0
: SEQ ID NO: 5
: LENGTH: 1542
: TYPE: DNA
: ORGANISM: Human
: OS-09-724-519-5

```

Query Match	1.0%;	Score 95.8;	DB 4;	Length 1542;
Best Local Similarity	53.5%;	Pred. No. 5.8e-14;		
Matches 286; Conservative	0;	Mismatches 222;	Indels 27;	Gaps 3

QY	684	AAAAGGTGAAAAAACAAGCATTATGAGAGAGCTAAAATGATATGCTACTACTGCTT	743
Db	644	AAAAGGGGACGAAAAAAGSACAACGTGACGTACTGTATGATGATGCTACTACTGCTT	703
QY	744	CACATACAAATTTAGAAATGATTTGTTGAAAGCCGAGACAAAAATGATCCCAAAATTAG	803
Db	704	CCACCTAGATTTCTCTGTTCACATACATATGAAAA-----AACTAGCA	748
QY	804	AGAACTGTATGAGAGCTGTCATGATGATACATGAAATTTGGTAGATCTGTGGCAGTG	863
Db	749	TTGATGAGAGAAGAGCTGTTTAAATGCGAAAGTGAACCTGGTTGATCTTCGAGAAAGTG	808
QY	864	AAAGAGCAAGCCAAACTGGAGCTGAAAGGTGTGAGACTTAAGSAGAGGCTGCAACATCAAC	923
Db	809	AAAACATATGGCCCTTCTGGAGCTGTGGAATAGAGAGAGCTGGAGACCTGGAAATATATATC	868
QY	924	GCAGCTGTATTCCTGTGGACAGGTATTAAGAAGCTGTGACAGCGCCAGGTGGTGGAT	983
Db	869	AATCCCTGTGACTTTGGGAAGGGTATTAATTCGCTCTTGTAAAGAAACACCTCATG---	925
QY	984	TTATTAACCTACAGACAGCAAACTCAGCAGAAATTTCTCCAAAATTCATTGGAGAGAAATG	1043
Db	926	---TTCCCTTATCCAGATCTTAACCTAATCTAGATCCCTCAGAGATTTCTGTGGAGGCGCTA	982
QY	1044	CTAAACGGTATTAATTTGGACAATTAAGCCAGATTG-----TTTATGTAGACTCTAA	1097
Db	983	CAGAGACATCTAATTAATTTGCACAACATTTCTCTGCACTCTGCAATCTTGAGAACTCTGGA	1042
QY	1098	GTAACACTTCAGTTTGGCAGTACTGCCAAACATGTGAGAAATTAATCTCCCATGTTAATGAG	1157
Db	1043	GTACATTTGAAATATGCTCATAGAGCAAAAGAACATATTGAATTAAGCCTGAAGTAACTGAGA	1102
QY	1158	TTCTTGATATGAAGCGTTGCTAAAAAGGTACAGAAAGAAATCTTGGATTAAAA	1212
Db	1103	AATCCACCAAAAAAGCTCTTATTAAGAGATTAATCGAGAGATTAAGACGTTTAAAA	1157

```

RESULT 14
US-09-592-037-5
: Sequence 5, Application US/09592037
: Patent No. 6437115
: GENERAL INFORMATION:
: APPLICANT: Wood, Kenneth
: APPLICANT: Finer, Jeffrey
: APPLICANT: Berard, Christophe
: APPLICANT: Mak, John
: APPLICANT: Sakowicz, Roman
: TITLE OF INVENTION: Methods of screening for modulators of
: TITLE OF INVENTION: cell proliferation and methods of diagnosing cell
: TITLE OF INVENTION: proliferation states
: FILE REFERENCE: 1014A
: CURRENT APPLICATION NUMBER: US/09/592,037
: CURRENT FILING DATE: 2000-06-12
: PRIOR APPLICATION NUMBER: 09/428,156
: PRIOR FILING DATE: 1999-10-27
: NUMBER OF SEQ ID NOS: 10
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 5
: LENGTH: 1542
: TYPE: DNA
: ORGANISM: Human

```

US-09-592-037-5

```

Query Match      1.0%; Score 95.8; DB 4; Length 1542;
Best Local Similarity 53.5%; Pred. No. 5.8e-14;
Matches 286; Conservative 0; Mismatches 222; Indels 27; Gaps 3,

```

QY	684	AAAGGGGTGAAAAAACAAGACATTATGAGAGACTAAATGATGATGATCAATGAGTCCGTT	743
Db	644	AAAAAGGGGGCAGCAAAAAAGCAACACACACACTACTCTGATGAAATGCATCTAGTCGTT	703
QY	744	CACATACAAATATTTAGATGATGTTGGTAAAGCCGACAGACAAGAAATGATCCACAAATTCAG	803
Db	704	CCCCTCAGATTTCTCTGTTACAAATACATATGAAAGT-----AACTACGA	748
QY	804	AGACCTGATGAGAGCGTGCATGATGATCTACATGAAATTTGGTATCTTGGTGGCAGAG	863
Db	749	TTTGATGAGAGAGACCTGTTTAAATCGGAAAGTTGAATCTGGTGTACTTGCAGGAAGTG	808
QY	864	AAAGAGCAAGCCAAACTGAGCTGAAAGCTGAGACTTAAAGAAAGGCTCAACATCAACC	923
Db	809	AAAACATTTGGCCCTCTGAGAGCTGTTGATTAAGAGAGCTGGGAAGCTGGAAATATTAATC	868
QY	924	GCACCTGTTTATTCCTTGGAGACGATTTAAGAAGCTTTACGCGAGCCAGCGCTGGTGAT	983
Db	869	AATCCCTGTGTGACTTTGGGAAGGCTCTTACTGCCCTGTGAGAAACAACCTCATAG---	925
QY	984	TTATTAACCTACAGAGACAGCAAACTCACCGAATTTCCAAAATTTGATGGAGGAATG	1043
Db	926	---TTCTTATACGAGATCTAAACTAATCACTAGAAATCCCGAGATTTCTTTGGAGGGCGTA	982
QY	1044	CTAATAAGGTTATTAATTTGCACATTTACGCCAGTTTC-----TTTGTATGAGACTCTAA	1099
Db	983	CAAAACAATCTATATATGCAACAATTTCTCTCATCTCTCAATCTTTGGAGAAACTCTGA	1043
QY	1098	GTACACTTCAGTTTGGCCAGTACTGCCAAACATGTGAAATTAATCTCCCATGTTAATGAG	1157
Db	1043	GTACATTTGGAATATGTCATATGAGCAAAAGAAATATTTGATTAATAGCCTGAAGTAATCNA	1100
QY	1158	TCTCGATGATGAAGCGTGTCTAAAAAAGTACAGAAAGAAATCTTGATTTTAA	1212
Db	1103	AACGTACCAAAAAAGCTCTTATTTAAGAGATATACGAGAGGATAGAAAGCTTTAAA	1157

```

RESULT 15
US-09-724-519-7
Sequence 7, Application US/09724519
Patent No. 6414121
GENERAL INFORMATION:
APPLICANT: Wood, Kenneth
APPLICANT: Finer, Jeffrey
APPLICANT: Beraud, Christophe
APPLICANT: Mak, John
APPLICANT: Sakowicz, Roman
TITLE OF INVENTION: Methods of screening for modulators of
TITLE OF INVENTION: cell proliferation and methods of diagnosing cell
TITLE OF INVENTION: proliferation states
FILE REFERENCE: 1014A
CURRENT APPLICATION NUMBER: US/09/724,519
CURRENT FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 09/592,037
PRIOR FILING DATE: 2000-06-12
PRIOR APPLICATION NUMBER: 09/428,156
PRIOR FILING DATE: 1999-10-27
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 1728
TYPE: DNA
ORGANISM: Human
US-09-724-519-7
Query Match 1.0%; Score 95.8; DB 4; Length 1728;
Best Local Similarity 53.5%; Pred. No. 6.1e-14;

```


Search completed: May 23, 2003, 01:06:49
Job time : 482 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 23, 2003, 00:54:39 ; Search time 1088 Seconds
(without alignments)
11682.693 Million cell updates/sec

Title: US-09-150-867-2

Perfect score: 9626

Sequence: 1 gaattccgagtcgtagtag.....ttaaaaaaaaaaagcaattc 9626

Scoring table:

Gapop 10.0 , Gapext 1.0

Searched: 828747 seqs, 660231138 residues

Total number of hits satisfying chosen parameters: 1657494

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCR_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCITS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	620.4	6.4	8493	9 US-10-071-766-51	Sequence 51, App1
2	615.2	6.4	7992	9 US-09-893-519A-140	Sequence 140, App
3	206.6	2.1	581	10 US-09-833-790-151	Sequence 351, App
4	138.2	1.4	501	10 US-09-833-790-167	Sequence 167, App
5	117.6	1.2	3366	9 US-09-938-842A-2651	Sequence 2651, App
6	95.8	1.0	4858	10 US-09-954-456-733	Sequence 733, App
7	90	0.9	29793	10 US-09-973-451-38	Sequence 38, App1
8	85.6	0.9	501	10 US-09-833-790-329	Sequence 329, App
9	81.2	0.8	6586	10 US-09-954-456-1115	Sequence 1115, App
10	81.2	0.8	6586	10 US-09-954-456-1790	Sequence 1790, App
11	73.8	0.8	3348	10 US-09-954-456-90	Sequence 90, App1
12	73.8	0.8	3348	10 US-09-954-456-727	Sequence 727, App
13	68.2	0.7	5361	9 US-10-071-766-80	Sequence 80, App1
14	68	0.7	5361	9 US-09-742-096-2	Sequence 2, App1
15	68	0.7	6152	9 US-09-742-096-1	Sequence 1, App1
16	67.8	0.7	5848	12 US-10-044-090-299	Sequence 299, App
17	67.6	0.7	3930	10 US-08-847-874A-2	Sequence 2, App1
18	59	0.6	9539	9 US-10-239-676-52	Sequence 52, App1
19	58.4	0.6	1002	10 US-09-879-536-442	Sequence 442, App

20	58.2	0.6	498	9 US-09-854-133-345	Sequence 345, App
21	58.2	0.6	498	10 US-09-738-973-345	Sequence 345, App
22	58	0.6	529	10 US-09-983-965-2109	Sequence 2109, App
23	57.4	0.6	446	10 US-09-960-352-3400	Sequence 3400, App
24	55.6	0.6	1267	12 US-10-001-843-45	Sequence 45, App1
25	54.4	0.6	451	10 US-09-960-352-10262	Sequence 10262, App
26	53.2	0.6	728	10 US-09-764-898-109	Sequence 109, App
27	53.2	0.6	1488	10 US-09-764-898-37	Sequence 37, App1
28	53	0.6	516	10 US-09-960-352-5785	Sequence 5785, App
29	52.8	0.5	589	10 US-09-864-721-4741	Sequence 14741, App
30	50.6	0.5	12405	9 US-10-239-676-35	Sequence 35, App1
31	50.4	0.5	554	9 US-10-101-487-69	Sequence 69, App1
32	50.4	0.5	554	9 US-10-101-487-106	Sequence 106, App1
33	50.4	0.5	1014	10 US-09-883-096-3	Sequence 3, App1
34	50.4	0.5	1152	10 US-09-883-096-6	Sequence 6, App1
35	50.4	0.5	4108	10 US-09-883-096-1	Sequence 1, App1
36	50.2	0.5	522	9 US-10-101-487-71	Sequence 71, App1
37	50.2	0.5	530	9 US-10-101-487-73	Sequence 73, App1
38	49.4	0.5	7657	9 US-10-239-676-185	Sequence 185, App1
39	49	0.5	17848	9 US-09-938-842A-646	Sequence 646, App
40	48.8	0.5	2241	9 US-10-171-311-7	Sequence 7, App1
41	48.6	0.5	12313	9 US-10-171-311-7	Sequence 7, App1
42	48.6	0.5	12337	9 US-10-171-311-5	Sequence 5, App1
43	48.6	0.5	12438	9 US-10-171-311-3	Sequence 3, App1
44	48.6	0.5	12462	9 US-10-171-311-1	Sequence 1, App1
45	48.4	0.5	431	10 US-09-960-352-5558	Sequence 5558, App

ALIGNMENTS

RESULT 1
US-10-071-766-51
Sequence 51, Application US/10071766
Publication No. US20020192678A1
GENERAL INFORMATION:
APPLICANT: Huel-Mei Chen
TITLE OF INVENTION: GENES EXPRESSED IN SENESENCE
FILE REFERENCE: PA-0043 US
CURRENT APPLICATION NUMBER: US/10/071,766
CURRENT FILING DATE: 2002-02-07
NUMBER OF SEQ ID NOS: 144
SOFTWARE: PERL Program
SEQ ID NO 51
LENGTH: 8493
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No. US20020192678A1 441283.3
US-10-071-766-51

Query Match 6.4%; Score 620.4; DB 9; Length 8493;
Best Local Similarity 65.5%; Pred. No. 9e-132;
Matches 1001; Conservative 0; Mismatches 491; Indels 36; Gaps 5;

QY	69	GGCGCAAAAGGAGCACTAAAGTGACAGAGAGGAGCGGTGCGTACCGATTTCGCAC	128
DB	17	GGGGCTGTGAGCCCTGAAAGTCCCGCCGCGGAGGGTCTCTGCGCATTTTGTGGACACG	76
QY	129	TATCGGTCTCAAAATGTCCGAGGAGATGACATTAAATGTGTGTGAGGTTTGGCCGC	188
DB	77	TTACAGCTATATGATGAGGCGGAGGAAGAGCGGTCTGCGTGCAGTGGCGCCGC	136
QY	189	TTATACAGAGACACAA-----GGGATCAAGCCACCTGCATGAGAGGCTGCAACA	242
DB	137	TCAAAGCAGACAGATCACTTGGGAAACCTGCCAATTACGGAACACTGCACATA	196
QY	243	ACACATTTCCCAAGTGTGAGCAAAAGTCTTCAATTTGATGTGATTAATTC	302
DB	197	ATGCTATTATTAAGTTGATGAGTAAGTAATTCCTTCATTTTATGCTGCTTCATGTA	256
QY	303	ACGAATCAACAAGTCAAAATTACCAAGAAATAGCAGTACTATCATCATCAGCTTTC	362

```

Db 257 ATGAACATCAAAATGATGATGAAGAAATAGACACCAATCATGATCTGGCACTAC 316
QY 363 AGGATATATGCGCAATATTTGGATACGACACATCTTCGGCGAAGACATACAA 422
Db 317 AAGGCTACAAATGATGATATTTGGCTATGACACAGATGCTTTCGAGAAACATATAC 376
QY 423 TGATGGGAACCAAAATTCATTTGGCATATATCCCAAGCCATACAGAGATTTTAA 482
Db 377 TGATGGGTTTCAAGATCATTTGGAGATTATCCAGGGCAATTCAGACATTTCCAA 436
QY 483 TTATTCAGGATACCCAGACAGAGATTTCTTCAAGATTTCTTATATGAGATTTCA 542
Db 437 AAATTAAGAAATTCCTGATAGGAAATTTCTTCAAGATTTCTTATATGAAATATACA 496
QY 543 ATGAACATGTCGAAAGACCTATGATGATGACAGAAAGAAAGACCCCTGGAAATGGCG 602
Db 497 ATGAACCATATACAGATTTTCTGTCGCACTCAAAATTAACCTTATTTTGGAG 556
QY 603 AGGATTTTATAGAAACGATGATTTGCTGACCTGATGAGAACTTGAATGCTCTG 662
Db 557 AAGATGTCATATGAAATGATGATGATGCTGATCTACAGAAAGATGATATACATCAG 616
QY 663 AACATGTAATACATGATGATCAAAAGGATGAAACAAACGATATGAGAGCTAAA 722
Db 617 AATAGCTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 676
QY 723 TGAATGATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 782
Db 677 TGAATCAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 736
QY 783 GAAATGATCCCAAAATTCAGACACATGATGATGATGATGATGATGATGATGATGAT 842
Db 737 AGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 787
QY 843 TGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 902
Db 788 TGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 847
QY 903 AGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 962
Db 848 AGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 907
QY 963 GCGAGGCGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1022
Db 908 GCGAGGCGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 967
QY 1023 AAAATTCATGAGAGAAATGCTAAAACGCTTATATTTGACAAATTCAGGATTTCT 1082
Db 968 AGAATTCCTTGGGAGGAAATCCAAAGACGATTTATCTGACAAATTTCTGATCTT 1027
QY 1083 TTGATGAGACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1142
Db 1028 TTGATGAACTCTTACCTGCTCCAGTTGCTCCAGTACTGCTAAATATATGAAATCTC 1087
QY 1143 CCCATGTTATGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1202
Db 1088 CTTATGTTATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1147
QY 1203 TGGATTTTAAAGAAATTCATGAGATGATGATGATGATGATGATGATGATGATGATGAT 1262
Db 1148 TGGATTTTAAAGAAATTCATGAGATGATGATGATGATGATGATGATGATGATGATGAT 1201
QY 1263 TGGCTTAAAGAGACATACACAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1322
Db 1202 TGGAAAGAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1261
QY 1323 AAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1379
Db 1262 ATGAGAAATTTGAAATCTTACAGGATGATGATGATGATGATGATGATGATGATGATGAT 1321
QY 1380 AGGACCAAGGATCAAGCAAGAAAGATGATGATGATGATGATGATGATGATGATGATGAT 1439

```

```

Db 1322 AGAATTAAGAGCTAAAGAAACGAAGATTTACTGTTGCTTGGCAAAATTAACAAA 1381
QY 1440 GTTTCATGCTTCTGGGTTTCTGATGATGATGATGATGATGATGATGATGATGATGAT 1499
Db 1382 TGAAGAA-----CTCAACATGATGATGATGATGATGATGATGATGATGATGATGAT 1429
QY 1500 GCAAGAGGCAAGATGCTGACATGATGATGATGATGATGATGATGATGATGATGATGAT 1559
Db 1430 CAACAAAACACATACATCTTCTATATATTTATACAGAAATGATGATGATGATGATGAT 1489
QY 1560 CAGAGTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1587
Db 1490 CAGAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1517

RESULT 2
US-09-893-519A-140
; Sequence 140, Application US/09893519A
; Publication No. US20030027243A1
; GENERAL INFORMATION:
; APPLICANT: ANADYS PHARMACEUTICALS, INC.
; APPLICANT: THOMPSON, Craig
; APPLICANT: MOORE, Jeffrey
; APPLICANT: BURMAN, Ed T.
; APPLICANT: BRADLEY, John
; APPLICANT: DESILVA, Thimara
; APPLICANT: HARRIS, Sandra
; APPLICANT: KOMARITSKY, Svetlana
; APPLICANT: MENDILLO, Marc
; APPLICANT: MOORE, Daniel
; APPLICANT: MCCOY, Melissa
; APPLICANT: SANDERSON, Karen
; APPLICANT: HAO, Tareq
; APPLICANT: ZHU, Shuhao
; APPLICANT: LONG, Fan
; APPLICANT: DAVIDOV, Eugene
; TITLE OF INVENTION: ANTIFUNGAL COMPOUNDS AND METHODS OF USE
; FILE REFERENCE: 0342/16548-US2
; CURRENT APPLICATION NUMBER: US/09/893,519A
; PRIOR FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 60/215,164
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: US 60/224,457
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 146
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 140
; LENGTH: 7992
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: Human Genbank/215005.1
; DATABASE ENTRY DATE: 1993-01-10
; RELEVANT RESIDUES: (1)..(7992)
US-09-893-519A-140

Query Match 6.4%; Score 615.2; DB 9; Length 7992;
Best Local Similarity 66.7%; Pred. No. 1.4e-130;
Matches 970; Conservative 0; Mismatches 448; Indels 36; Gaps 5;

```

```

Db 181 AATGTATGAAGAAATACAGACCAATCATGATTCGGCATACAAAGCTACATGCT 240
Qy 377 ACAATATTTGCTACGAGACAGACATCTTACGACAGACGTACACATGATGGACACCA 436
Db 241 ACTATATTTGCTACGAGACAGACATCTTACGAGACAGACATGATGGATGCTACGA 300
Qy 437 AATTCATTTGGCATATATCCCAAGCATACAGAAATTTTAAATTTTTCAGAGATA 496
Db 301 GATCATTTGGGAGTATATCCAGGCAATTCATGACATTTCCAAAAATTAAGAAATTT 360
Qy 497 CCGAAGAGAGAGATCTTCTTAAGAGTTTCTTATATGAGATTTACATGAAACCTGTGAA 556
Db 361 CCTGATAGGAGATTTCTCTTACGTGATCTTACATGGAATTTACATGAAACCATTACA 420
Qy 557 GACCTACTGTGTGATGACAGAAAGAAAGCCCTTGAAATTCGCGAGATTTTAAATGA 616
Db 421 GATTTACTCTGTGGACATCCAAAAATGAAACCTTTAATTTTCGAGAAAGTCTCATAG 480
Qy 617 AACGTGTATGTGCTGACCTGACCTGAGAACTTGTAACTGTTCTGCAACATGTAATACG 676
Db 481 AATGTGTATGTGCTGATCTCAGAGAGAGATTTATATACATGAAATGCTTTGAAA 540
Qy 677 TGGATCAAAAAAGGCTGAAAAAAGACATTTATGAGAGACTAAATGATGATCATAGT 736
Db 541 TGGATTTACAAAGGAGAAAAAGACAGGCAATTATGAGAAACAAAAATGATACAAAGAAC 600
Qy 737 ACTGCTTCAATCAATATTTTGAATGATTTGTTGAAAGCCGAGACAGAAATGATCCACA 796
Db 601 AGCTCTTCCATCATCATCTTTAGATGATTTTGGAAAGCAGAGAAAGGCTACACCTTCT 660
Qy 797 AATTCAGAAAGTGTGATGAGCTGTGATGATGATGATGATGATGATGATGATGATGAT 856
Db 661 AAT-----TGTAGAGATCTGTTAGATTCATTTGATTTGATGATGATGATGAT 711
Qy 857 GGCAGTGAAGAGCAAGCAAGCAAGCTGAGCTGAGAGTGTGAGACTTAAAGAGGCTGCAAC 916
Db 712 GGCAGTGAAGAGCAAGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGAGGCTGATAT 771
Qy 917 ATCAACCCGAGCTGTTTATCTGTTGAGACAGTTTATTAAGAGCTTAAAGCGCCAGCT 976
Db 772 ATAAATCGAAGCTTATTTATTTGAGACAAAGTATCAAAACTTATGATGAGCAAGT 831
Qy 977 GGTGATTTTAACTATACAGAGACAAAGTCAACAGATTTCTCCAAATTCATTTGAGGA 1036
Db 832 GGTGATTTTAACTATATGAGATGAGATTAACAGATTTCTGAGATTTCTGAGGA 891
Qy 1037 GGAATGCTAAAAAGGTTATATTTGCAATTTAGCCAGTTTCTTTGATGAGACTTA 1096
Db 892 GGAATTCAGAAAGACAGATATATGCAATTTAGCCAGTATCTTTGATGAGAACTTT 951
Qy 1097 AGTACACTTCACTTCCAGCTACTGCAAAACATGTGAGAAATTTCTCCATGTTATGAG 1156
Db 952 ACTGCTTCCACTTTGCAAGTACTGCTAAATATGAAATACTCTTATTTATGAG 1011
Qy 1157 GTCCTGATGATGAGAGCGTGTGTTAAAAAGGTACAGAAAGGAATGTTGATTTAAAGAA 1216
Db 1012 GATCAACTGATGAGAGCTTCTGTAAGAGGTATGAAAGAAATTAATGATCTTTAAAAA 1071
Qy 1217 CAATTAGAAATTTAGAGTCACTGCTGAAACAAAGCTCAAGCAATGCTTAAAGAGAG 1276
Db 1072 CAATTAGAG-----GAGGTTCTTTAGAGAGCGGGGCTCAAGCAATGAGAAAAAGCCA 1125
Qy 1277 CATACAGAGTGTGCTAGCTGAATCAACACATACAGAAAGAGAGAGATGAAATATG 1336
Db 1126 TTGGCCCAACTTTTGAAGAAAGATTTTCAAGAAAGTACAGAAATGAGAAATTTGAA 1185
Qy 1337 CACTGACAAATATTTGTTGT---TGCTTCATCCCAAGATTTCAACAGGACCAAAAGGCG 1393
Db 1186 AACTTAAACAGGATCTGTGACCTCTTCTTCTTCCCTCAGCTTCAACAGGAAATTTAAAGGCT 1245
Qy 1394 AAACGAAAAAGAGATTAAGTGGGCGCAGGAAAAATCCAAAAATAGTTTACATGCTTCT 1453

```

```

Db 1246 AAAAGAAAAAGAGATTAAGTACTGTGCTTGGCAAAATTAACAAATTAAGAAAGCTCAAC 1305
Qy 1454 GGTGTTTCTGACTTTATATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTAT 1513
Db 1306 TATGAGATCAATTTATATATACCAACA-----AATTAACAAACAAACAAACAAAT 1353
Qy 1514 TTCTGACATGCTTCCAGAAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1573
Db 1354 AAGCTTCTTATTAATTTATATATGAGAAATGATGATGATGATGATGATGATGATGATGAT 1413
Qy 1574 TTTGATGAGCCCT 1587
Db 1414 TTCAGTAACATCT 1427

RESULT 3
US-09-833-790-351
; Sequence 351, Application US/09833790
; Patent No. US20020068288A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Wang, Tonglong
; APPLICANT: Secret, Heather
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Indrias, Carol Y.
; APPLICANT: Pan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.512
; CURRENT APPLICATION NUMBER: US/09/833,790
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 351
; LENGTH: 581
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-790-351

Query Match 2.1%; Score 206.6; DB 10; Length 581;
Best Local Similarity 68.4%; Pred. No. 2.6e-37;
Matches 333; Conservative 0; Mismatches 144; Indels 10; Gaps 3;

Qy 955 GAAGCTTACGAGCGGCGGAGCTGTGCTGATTTAACTACAGAGACAGAACTCACAG 1014
Db 72 GAAGCTTACGAGCGGAGCTGTGCTGATTTAACTACAGAGACAGAACTCACAG 131
Qy 1015 AATTCCTCAAAATTTCAATTTGGAGAAATGCTAAACGTTATTAATTTGCAATTTACGCC 1074
Db 132 AATTCCTCAAAATTTCAATTTGGAGAAATGCTAAACGTTATTAATTTGCAATTTACGCC 191
Qy 1075 AGTTCTTTTATGAGAGCTCTAATGACCTTCACTTCACTTCACTTCACTTCACTTCACTT 1134
Db 192 AGTTCTTTTATGAGAGAGCTCTAATGACCTTCACTTCACTTCACTTCACTTCACTTCACTT 251
Qy 1135 AATATCTCCCATTTAATGAGAGCTGCTGATATGAGAGCTGCTAATTAAGTACAGAA 1194
Db 252 GAATATCTCTTATGATTAAGAGATCAACTGATGAGAGCTGCTGATTAAGTACAGAA 311
Qy 1195 GGAATCTTGGATTTTAAAGAAACATTTAGAGATTTAGAGTATGCTGTAAGAAAGC 1254
Db 312 AGAATATTAATGATCTTAAAGAAACATTTAGAG-----GAGGTTCTTTAGAGAGCGGGCG 365
Qy 1255 TCAAGCAATGCTTAAAGAGAGATACACAGTGTGCTAGTGAATCAAAACATACACA 1314
Db 366 TCAGGCAATGAGAAAGAACCAATTTGGCCCACTT-TTGGAAAGAAAGATTTCTTCAGAA 424
Qy 1315 AGAGAGAGAGATTAAGATTAAGCACTGACCTGACAAATATGTTGT---TGCTTCATCCCAAA 1371
Db 425 AGTACGAATGAGAAATTTGAAACCTTAACAGAGATGCTGTGACCTCTTCTTCCCTCAC 484
Qy 1372 ATCTCAACAGCAAAAGGCTCAACGAAAGAGATTTAGTGGGCGCAGGAAAAAT 1431

```

Db 485 GTTGCAACGAGAAATTAAGGCTTAAGAAAAAGAGAGATTACTTGCTGCTTGCATAAAT 544
 QY 1432 CCAAAAT 1438
 Db 545 ACCAAAT 551

RESULT 4
 US-09-833-790-167

; Sequence 167, Application US/09833790
 ; Patent No. US20020068288A1

; GENERAL INFORMATION:

; APPLICANT: Lodes, Michael J.

; APPLICANT: Wang, Tonglong

; APPLICANT: Secret, Heather

; APPLICANT: Mohamath, Raedoh

; APPLICANT: Indrias, Carol Y.

; APPLICANT: Fan, Liqun

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE REFERENCE: 210121.512

; CURRENT APPLICATION NUMBER: US/09/833,790

; NUMBER OF SEQ ID NOS: 440

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 167

; LENGTH: 501

; TYPE: DNA

; ORGANISM: Homo sapien

US-09-833-790-167

Query Match 1.48; Score 138.2; DB 10; Length 501;
 Best Local Similarity 57.78; Pred. No. 1.3e-21;

Matches 266; Conservative 0; Mismatches 193; Indels 2; Gaps 1;

QY 2844 AATGGAAGCTGATTTGCAAGAAAGATTGCAAGTCTTTAATGAGATTACTTAA 2903
 Db 1 AATGGAAGCTGATTTGCAAGAAAGATTGCAAGTCTTTAATGAGATTACTTAA 2903

QY 2904 ATGGCTCTGCGAGAGAAAGTCCCAAGATTGCTTTCGCTGATTAAGTAA 2963
 Db 61 CCTCCTTATGAGTGGCAAGTCTCCAAAGATTGCTCTGTAATTTGGAATTTGGAAGAA 120

QY 2964 AGTTTCTGATTTCAAGAGAGCTTGAAGACA--TTGGAAGAAAAAATGCTTGA 3021
 Db 121 AGATTACTGATCTTCAGAAAGAACTAATTAAGAAAGTTGAAGAAAAATGAGCTTTGCG 180

QY 3022 GAATGAGTCTGCTTCAAGTAATTAATTTTGGCAATGAGTTGATCTTGA 3081
 Db 181 GGAAGAACTGATTTTCTTCAAGTAATTAATTTTACCTTCTGAAGTGAAGAGCTGAG 240

QY 3082 AATCAGATCGCAAGGCTCTGAAGAGATTAATTTAAGCAAGAGAGAACTTC 3141
 Db 241 GAAAGAGTACCAACAAATCTGAAGAGCTCCATATTAATCAACAAAAAGTAAATT 300

QY 3142 TGCATCTATTAAAGCAACAAAGATTATCATGACAGAGCAATCTGAGCATTTTACA 3201
 Db 301 GTTTCTGAAGTACTCATTAAGAGAGTAAAGTTTCAAGTTTACTTTAAGAAATTTGGAA 360

QY 3202 ACTGATCGAGAGTACACACACACAGTCAAAAAGTGCAGACAGCTGAGAGCAATCTT 3261
 Db 361 AACAAAGAGTACCTTACACTACAGTCAAGTCAATTAATTAAGCACTGATCAAGAAATTC 420

QY 3262 GGAGATGAGAAAAATGATGATCTTTTGAAGTATA 3302
 Db 421 AATTTCAAAACCTTATGAGCTTTGAGCAAAAGTATA 461

RESULT 5
 US-09-938-842A-2651

; Sequence 2651, Application US/09938842A
 ; Patent No. US20020160378A1
 ; GENERAL INFORMATION:

; APPLICANT: Harper, Jeff
 ; APPLICANT: Kreps, Joel
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; FILE REFERENCE: SCRIPI300-3
 ; CURRENT APPLICATION NUMBER: US/09/938,842A
 ; PRIOR FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: US 60/227,866
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: US 60/264,647
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/300,111
 ; PRIOR FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 5379
 ; SEQ ID NO 2651
 ; LENGTH: 3366
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-09-938-842A-2651

Query Match 1.28; Score 117.6; DB 9; Length 3366;
 Best Local Similarity 54.28; Pred. No. 2.3e-16;
 Matches 412; Conservative 0; Mismatches 309; Indels 39; Gaps 7;

QY 471 AAGTTTAAATTAATTCAGAGATACCGAAGAGATTCTCTAAGATTCTTATA 530
 Db 830 AGATGTAAGTGTCTTACCGGAGCTCAGAGAGGAGTTCTCTCGTGTTCATATC 889

QY 531 TGAAGATTACATGAAGAACTGTGAAGACCTACTGTGTGATGACGAGAAAGAGCCCT 590
 Db 890 TTGAATATCAATGAAGAGTGTAAATGATTACTG-----GATCCAAAGGCCAAACT 943

QY 591 TGAAGTTCCGAGGATTTAATGAAGAGTGTGTGTGCTGACCTGACGAAAGACTTG 650
 Db 944 TACGTGTAGAGAGATTTCCAGGG---CACTATGTTGAGGTATCAAGAGAAAGTTG 1000

QY 651 TAAATGTTCTGCAATGATATACAGTGTATCAAAAAGGTGAAAAAAGACATTAATG 710
 Db 1001 TTTGTCTCCGCGCATGACATATCATTTGAGCTGGGGAAGAAACATGCTCATGTTG 1060

QY 711 GAGGACTAAATGAATGATCATATGATGTGTTTCACTATCAATATTTAGATGATTTG 770
 Db 1061 GTTCAAAATTAATCAATTTGTTGAGCAGAGTCACTAATTTTACACTGATGTTG 1120

QY 771 AAGCCGAGACAGAAATGATCCCAAAATCAGAGAACTGTGAGAGCTGATGTAAT 830
 Db 1121 AAGTATG---CTACTGAGATGATATGATGA---GTTATCTTTT 1162

QY 831 CTCACCTGAATTTGATGATCTTCTGCGAGTGAAGAGCAACCAACTGAGCTGAAG 890
 Db 1163 CTCACCTGAATTTGATGATGATCTTCTGCGAGTGAAGAGCTGAGCTGAAGCTGAAG 1219

QY 891 GTGTGAGCTTGAAGAGGTGCAATCAACCCGAGCTTGTATTCCTTGGACAGTTA 950
 Db 1220 GATTGAGAGAGAGAGGCTTCATACATCAACAAAGGCTTCTTAACTCTTGGAACTGTGA 1279

QY 951 TTAAGAGCTTTAGCGAGCGGCTGTGATTTTAATTAAGTACAGACAGCAAACTGA 1010
 Db 1280 TTGGAAGACTTATGAGAGGTAAGAGCACTACATTTCA---TATGCTGACTTAAGCTTA 1336

QY 1011 CCAAGATTTCCAAATTAATTTGAGAGAAATGCTAAAGGTTATTAATTTGACAAATTA 1070
 Db 1337 CTCGTCTGTGCAATCTTCAATTAAGTGAAGAGATGATGCTGCTCATATGATTAATTA 1396

QY 1071 CGCGAGTTTC-----TTTGAATGAGACTCTAAGTACCTTAAGTTTGGCACTATGCA 1124
 Db 1397 CTCCTGCTGTCAGCACTAGTGAAGAACTCATTAACATTAAGTTTGGCACTAGGCGCA 1456

QY 1125 AACATGTGAAATATCTCCCATGTTAATGAGTCTGATGATGAAGGCTGTAAAAA 1184
 Db 1457 AGAGTATAGAAATATATGCTTTCACCGAATCAGATTATATGATGAGAGATCATTAATTAAGA 1516

Db	1707	TCCTAGGAAGCTGACGACCACTCTCTGAAAGACCTTCGAAGGCTTCATACCTCTC	1766
Qy	381	TATTTGCATACGACGACACATCTTTCAGGCAAGACGTACACATGATGGGACACCAATTT	440
Db	1767	TTTTTGGTTAAGGTCACACAGTGGCTCTGGAAAAATACATACGATGAAGGGATTTAGTGAAG	1828
Qy	441	CATTGGGCATATATACCCCAAGCCATACAGGAAGTTTTTAAATTTATTCAGGATACCGA	500
Db	1827	AACCAAGGATATATCCCAAGATTTTGTGAGATCTTTTTCTCAAGTACCCAGAAAAA	1888
Qy	501	ACAGAGAGTT-----TCCTCTAAGAGTTCTTATATGAGATTTATCAATGAAGTGTGA	554
Db	1887	CCCAAGAGGCTACAGCTATACATGTAATGAATGAGCTTCTTGAAGATATATGAAAAAATTC	1948
Qy	555	AAGACCTACAGCTGTG-----ATGACAGAAAGAAAGCCCTTGGAAATTCGG	602
Db	1947	ACGACCTTCTGGTTTGTAAAGATGAAATATGGCGAGAAAGCAACCACTGACAGATGAGGG	2006
Qy	603	AGGATTTTAAATAGAAAGCTATGTATGTGCTGACCTGACTAAGAAACTTGTAAATGTTCTGTG	662
Db	2007	AACATCTGTTTATGACACCATATGTGTGAAGCACTGTCAATGAAACATGTCACTTCTTACG	2066
Qy	663	AACATCTGATATACAGTGCATCAAAAAGGGTGAAAAACAGACATTATGGAGAGACTAAA	722
Db	2067	CTGATATCCAGAGAGTTGGCTGAAATTTGGGAAATTAACAAAGACCTACTGCTCTACTGGTA	2128

Db 2127 TGAAGATMAAAAGTTCCCGATCTCACTTACAGTTTCACCCCTGTGATGACCCAGACCAAGA 2186
 QY 783 GAATGATCCACAAATTCAGAGAACTGTGATGAGAGCTGTCAATGGTATCTCACTTGAATT 842
 Db 2187 CAGATTGTGTGAAGGGGAGAACACGATCACAGA-----ATTAACAAGTCAATTAACC 2240
 QY 843 TGGTAGATCTTGCTGGCGAGTGAAGAGCAACCAACTGAGAGCTGAAGAGTGTGAGACTTA 902
 Db 2241 TAATAGATCTGGCAGGCGAGTGAAGGCGCTCTACGGCTCACACTTAATGAGATGAGACTTA 2300
 QY 903 AGAAGGCTGCACATCAACCGCAGCTTGTATTCCTTGGACAGATTATTAAGAAGCTTA 962
 Db 2301 AGGAAGGTGTGAGTATTAATTAAGTCTGTCTGAATCTTGGAAAAGTTATATCTGCACCTT 2360
 QY 963 GCGA-----CGGCCAGGCTGGTGGATTATTAATCTACAGAGACAGCAAACTCACACAGA 1016
 Db 2361 CGGAACAAGCAAAACCAAGGAGTGTTTTATTCCTTAACGGAATCTGTTCATATGCG 2420
 QY 1017 TTCCTCAAAATTCATTTGGAGGAAATGCTAAACGGTATATTAATTTGCACATTTACGCCAG 1076
 Db 2421 TGTTAAAGAAAGTCTGGGTGGAATTCAAAACTGCATATTTGCTACGATTTAGTCCG 2480
 QY 1077 TTTC-----TTTTGATGAGACTCTAAGTACACTTCAGTTTGGCAGTACTGCCAAACATG 1130
 Db 2481 CTCGCACGACATAGAAAGAAACATTAAACACACCTAGATATGCTAAACCAACCCGGTTTAA 2540
 QY 1131 TGAGAAATACCTCCCATGTAAATGAGCTCCGATGATGA 1170
 Db 2541 TAGTCACACTGTAAAGTAAATGAAGATATGAACGGCTA 2580

RESULT 11
 US-09-954-456-90
 ; Sequence 90, Application US/09954456
 ; Patent No. US20020115057A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Young, Paul
 ; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C

FILE REFERENCE: 689290-76
 CURRENT APPLICATION NUMBER: US/09/954,456
 CURRENT FILING-DATE: 2001-09-18
 PRIOR APPLICATION NUMBER: US/60/233,617
 PRIOR FILING DATE: 2000-09-18
 PRIOR APPLICATION NUMBER: US/60/234,052
 PRIOR FILING DATE: 2000-09-18

PRIOR FILING DATE: 2000-09-20
 PRIOR APPLICATION NUMBER: US/60/234,923
 PRIOR FILING DATE: 2000-09-25
 PRIOR APPLICATION NUMBER: US/60/235,134
 PRIOR FILING DATE: 2000-09-25
 PRIOR APPLICATION NUMBER: US/60/235,637
 PRIOR FILING DATE: 2000-09-26
 PRIOR APPLICATION NUMBER: US/60/235,638
 PRIOR FILING DATE: 2000-09-26
 PRIOR APPLICATION NUMBER: US/60/235,711
 PRIOR FILING DATE: 2000-09-27
 PRIOR APPLICATION NUMBER: US/60/235,720
 PRIOR FILING DATE: 2000-09-27
 PRIOR APPLICATION NUMBER: US/60/235,840
 PRIOR FILING DATE: 2000-09-27
 PRIOR APPLICATION NUMBER: US/60/235,863
 NUMBER OF SEQ ID NOS: 2276
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 90
 LENGTH: 3348
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-954-456-90

Query Match 0.8%; Score 73.8; DB 10; Length 3348;
 Best Local Similarity 50.9%; Pred. No. 2.7e-06;
 Matches 202; Conservative 0; Mismatches 192; Indels 3; Gaps 1;

OY 597 TTCCGAGAGATTTTAAATGAAAGCTGTATGCTGACCTGACCTGAAGAAGCTTGAATGG 656
 DB 815 TTCCGAGAGATTTTAAATGAAAGCTGTATGCTGACCTGACCTGAAGAAGCTTGAATGG 874
 OY 657 TTCCGAGAGATTTTAAATGAAAGCTGTATGCTGACCTGACCTGAAGAAGCTTGAATGG 716
 DB 875 TTCCGAGAGATTTTAAATGAAAGCTGTATGCTGACCTGACCTGAAGAAGCTTGAATGG 934
 OY 717 TTCCGAGAGATTTTAAATGAAAGCTGTATGCTGACCTGACCTGAAGAAGCTTGAATGG 773
 DB 935 TTCCGAGAGATTTTAAATGAAAGCTGTATGCTGACCTGACCTGAAGAAGCTTGAATGG 994
 OY 935 TTCCGAGAGATTTTAAATGAAAGCTGTATGCTGACCTGACCTGAAGAAGCTTGAATGG 994
 DB 774 TTCCGAGAGATTTTAAATGAAAGCTGTATGCTGACCTGACCTGAAGAAGCTTGAATGG 833
 OY 995 TTCCGAGAGATTTTAAATGAAAGCTGTATGCTGACCTGACCTGAAGAAGCTTGAATGG 1054
 DB 834 TTCCGAGAGATTTTAAATGAAAGCTGTATGCTGACCTGACCTGAAGAAGCTTGAATGG 893
 OY 834 TTCCGAGAGATTTTAAATGAAAGCTGTATGCTGACCTGACCTGAAGAAGCTTGAATGG 893
 DB 1055 TTCCGAGAGATTTTAAATGAAAGCTGTATGCTGACCTGACCTGAAGAAGCTTGAATGG 1114
 OY 894 TTCCGAGAGATTTTAAATGAAAGCTGTATGCTGACCTGACCTGAAGAAGCTTGAATGG 953
 DB 1115 TTCCGAGAGATTTTAAATGAAAGCTGTATGCTGACCTGACCTGAAGAAGCTTGAATGG 1174
 OY 954 TTCCGAGAGATTTTAAATGAAAGCTGTATGCTGACCTGACCTGAAGAAGCTTGAATGG 990
 DB 1175 TTCCGAGAGATTTTAAATGAAAGCTGTATGCTGACCTGACCTGAAGAAGCTTGAATGG 1211

RESULT 12
 US-09-954-456-727
 Sequence 727, Application US/09954456
 Patent No. US20020115057A1
 GENERAL INFORMATION:
 APPLICANT: Young, Paul
 TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
 TITLE OF INVENTION: Sets
 FILE REFERENCE: 689290-76
 CURRENT FILING DATE: 2001-09-18
 PRIOR APPLICATION NUMBER: US/60/233,617
 PRIOR FILING DATE: 2000-09-18
 PRIOR APPLICATION NUMBER: US/60/234,052
 PRIOR FILING DATE: 2000-09-20

PRIOR APPLICATION NUMBER: US/60/234,923
 PRIOR FILING DATE: 2000-09-25
 PRIOR APPLICATION NUMBER: US/60/235,134
 PRIOR FILING DATE: 2000-09-25
 PRIOR APPLICATION NUMBER: US/60/235,637
 PRIOR FILING DATE: 2000-09-26
 PRIOR APPLICATION NUMBER: US/60/235,638
 PRIOR FILING DATE: 2000-09-26
 PRIOR APPLICATION NUMBER: US/60/235,711
 PRIOR FILING DATE: 2000-09-27
 PRIOR APPLICATION NUMBER: US/60/235,720
 PRIOR FILING DATE: 2000-09-27
 PRIOR APPLICATION NUMBER: US/60/235,840
 PRIOR FILING DATE: 2000-09-27
 PRIOR APPLICATION NUMBER: US/60/235,863
 NUMBER OF SEQ ID NOS: 2276
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 727
 LENGTH: 3348
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-954-456-727

Query Match 0.8%; Score 73.8; DB 10; Length 3348;
 Best Local Similarity 50.9%; Pred. No. 2.7e-06;
 Matches 202; Conservative 0; Mismatches 192; Indels 3; Gaps 1;

OY 597 TTCCGAGAGATTTTAAATGAAAGCTGTATGCTGACCTGACCTGAAGAAGCTTGAATGG 656
 DB 815 TTCCGAGAGATTTTAAATGAAAGCTGTATGCTGACCTGACCTGAAGAAGCTTGAATGG 874
 OY 657 TTCCGAGAGATTTTAAATGAAAGCTGTATGCTGACCTGACCTGAAGAAGCTTGAATGG 716
 DB 875 TTCCGAGAGATTTTAAATGAAAGCTGTATGCTGACCTGACCTGAAGAAGCTTGAATGG 934
 OY 717 TTCCGAGAGATTTTAAATGAAAGCTGTATGCTGACCTGACCTGAAGAAGCTTGAATGG 773
 DB 935 TTCCGAGAGATTTTAAATGAAAGCTGTATGCTGACCTGACCTGAAGAAGCTTGAATGG 994
 OY 935 TTCCGAGAGATTTTAAATGAAAGCTGTATGCTGACCTGACCTGAAGAAGCTTGAATGG 994
 DB 774 TTCCGAGAGATTTTAAATGAAAGCTGTATGCTGACCTGACCTGAAGAAGCTTGAATGG 833
 OY 995 TTCCGAGAGATTTTAAATGAAAGCTGTATGCTGACCTGACCTGAAGAAGCTTGAATGG 1054
 DB 834 TTCCGAGAGATTTTAAATGAAAGCTGTATGCTGACCTGACCTGAAGAAGCTTGAATGG 893
 OY 834 TTCCGAGAGATTTTAAATGAAAGCTGTATGCTGACCTGACCTGAAGAAGCTTGAATGG 893
 DB 1055 TTCCGAGAGATTTTAAATGAAAGCTGTATGCTGACCTGACCTGAAGAAGCTTGAATGG 1114
 OY 894 TTCCGAGAGATTTTAAATGAAAGCTGTATGCTGACCTGACCTGAAGAAGCTTGAATGG 953
 DB 1115 TTCCGAGAGATTTTAAATGAAAGCTGTATGCTGACCTGACCTGAAGAAGCTTGAATGG 1174
 OY 954 TTCCGAGAGATTTTAAATGAAAGCTGTATGCTGACCTGACCTGAAGAAGCTTGAATGG 990
 DB 1175 TTCCGAGAGATTTTAAATGAAAGCTGTATGCTGACCTGACCTGAAGAAGCTTGAATGG 1211

RESULT 13
 US-10-071-766-80
 Sequence 80, Application US/10071766
 Publication No. US20020192678A1
 GENERAL INFORMATION:
 APPLICANT: Huel-Mei Chen
 TITLE OF INVENTION: GENES EXPRESSED IN SENESECE
 FILE REFERENCE: PA-0043 US
 CURRENT FILING DATE: 2002-02-07
 PRIOR APPLICATION NUMBER: US/10/071,766
 NUMBER OF SEQ ID NOS: 144
 SOFTWARE: PERL Program
 SEQ ID NO 80
 LENGTH: 3376
 TYPE: DNA
 ORGANISM: Homo sapiens

THIS PAGE BLANK (USPTO)